Identification and characterization of novel class switch recombination factors

DISSERTATION
zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)
Im Fach Biologie/Molekularbiologie

eingereicht an der
Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von
M.Sc. Verónica Delgado Benito

Präsidentin
der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

Prof. Dr. Bernhard Grimm

Gutachter/innen

1. Prof. Dr. Michela Di Virgilio
2. Prof. Dr. Achim Leutz
3. Prof. Dr. Chiara Romagnani

Tag der mündlichen Prüfung: 23.06.2020

Table of contents

ABSTRACT

Class Switch Recombination (CSR) is the mechanism responsible for antibody isotype differentiation in mature B-lymphocytes and it is required for the establishment of an efficient immune response. CSR is a deletional somatic reaction that replaces the gene encoding for the μ constant region (donor) of the immunoglobulin heavy chain (*Igh*) with one of the several downstream constant genes (acceptor) within the same locus. Consequently, a B cell switches from expressing an antibody molecule of the IgM class to one of the other isotypes (IgG, IgE or IgA), which have the same antigen specificity but different effector function. At the molecular level, CSR occurs between donor and acceptor repetitive DNA elements, the so-called switch (S) regions that precede each constant gene. Upon antigen encountering, B cells proliferate and undergo cytokine stimulation that induces germ-line transcription (GLT) at the *Igh* locus. This process leads to non-coding transcripts and the exposure of single strand stretches of DNA, which are the substrate for the B cell specific enzyme AID (activation-induced cytidine deaminase). AID enzymatic activity generates U:G mismatches that are eventually converted into DNA double-strand breaks (DSBs). DSBs in donor and acceptor S regions are then directly ligated by the classical non-homologous end-joining pathway (c-NHEJ), leading to antibody isotype differentiation. Structurally, the *Igh* locus comprises 250 kb and contains two enhancers, Eμ and the 3'RR (3' Regulatory Region). While Eμ is essential during early stages of B cell development, the 3'RR controls antibody isotype differentiation. Altogether, CSR is a complex physiological process that involves the formation and repair of DSBs through different molecular mechanisms that have not been fully elucidated yet. For example, it has been shown that the 3'RR is required for GLT and therefore CSR but how the enhancer activity is regulated is still unknown. Therefore, the aim of this study was to identify and characterize novel CSR factors that play a role in any of these mechanisms, contributing to the establishment of an efficient immune response. To do so, a functional loss-of-CSR screen was set-up and performed in the CH12 lymphoma B cell line that can undergo antibody isotype differentiation to IgA *in vitro* with high efficiency. Specifically, potential 53BP1 and RIF1 interactors were somatically targeted by CRISPR/Cas9, since these two factors play a key role during CSR. In this study, the chromatin reader ZMYND8 was found to regulate GLT and CSR by binding and modulating the transcriptional activity of the 3'RR at the *Igh* locus. Moreover, PDAP1 was independently identified as a novel factor required for efficient CSR by modulating *Aicda* (AID) mRNA levels. Conclusively, the results of this thesis contributed to further understand the processes regulating antibody isotype differentiation and B cell activity during an immune response.

ZUSAMMENFASSUNG

Als Klassenwechsel (class switch recombination, CSR) bezeichnet man den Mechanismus der Antikörper-Isotyp-Differenzierung in ausgereiften B-Lymphozyten. Er wird für den Aufbau einer effizienten Immunantwort benötigt. CSR ist eine löschende, somatische Reaktion, welche das Gen für die konstante Domäne μ (Spender) der schweren Kette des Immunoglobulins (immunoglobulin heavy chain, Igh) ersetzt. Dabei erfolgt der Austausch mit einem von mehreren, strangabwärts gelegenen Genen (Akzeptor) im selben Lokus. Dadurch wechselt eine B-Zelle von der Expression eines Antikörpermoleküls der IgM-Klasse zu einer der anderen Isotypen (IgG, IgE oder IgA). Diese haben dieselbe Spezifität, jedoch eine andere Effektorfunktion.

Auf molekularer Ebene findet der Klassenwechsel zwischen Spender und Akzeptor repetitiver DNS-Elemente statt. Diese befinden sich in der Switch-Region (S), welchen sämtlichen konstanten Genen vorangeht. Wenn B-Zellen auf Antikörper treffen, vermehren sie sich und es erfolgt eine Stimulation durch Zytokine, welches eine Transkription der Keimbahn (germline transcription, GLT) am *IgH* Lokus hervorruft. Dies führt zu nicht-kodierenden Transkripten und dem Offenlegen von Einzelsträngen der DNS. Diese sind die Bindestellen für das B-Zellen-spezifische Enzym AID (activation-induced cytidine deaminase).

AID Enzymaktivität erzeugt U: G Basenpaarfehler, welche zu DNS-Doppelstrangbrüche (DNA double-strand breaks, DSBs) umgewandelt werden. DSBs in Spender und Akzeptor S-Regionen werden direkt durch den klassischen non-homologous end-joining (c-NHEJ) Pfad repariert, welcher zur Differenzierung der Antikörper-Isotypen führt.

Strukturell besteht der *Igh*-Lokus aus 250 kB und enthält zwei Enhancer, Eμ und 3'RR (3' regulatorische Region). Während Eμ essentiell für die frühen Phasen der B-Zellenentwicklung, kontrolliert 3'RR die Differenzierung der Antikörper-Isotypen.

Zusammengenommen ist CSR ein komplexer, physiologischer Prozess, welcher die Formation und Reparatur von DSBs umfasst. Die verschiedenen molekularen Mechanismen, welche daran beteiligt sind, sind noch nicht vollständig verstanden. Beispielsweise wurde gezeigt, dass 3'RR benötigt wird für GLT und dadurch CSR. Wie diese Enhancer-Aktivität reguliert wird, ist jedoch noch unbekannt. Daher war das Ziel dieser Studie die Identifikation und Charakterisierung neuer CSR-Faktoren, welche diese Mechanismen beeinflussen und zum Aufbau einer effizienten Immunantwort beitragen.

Dafür wurde ein Testverfahren für einen funktionalen Verlust von CSR aufgebaut und durchgeführt in CH12 Lymphoma B-Zellen welche die Differenzierung von Antikörper-Isotypen nach IgA *in-vitro* mit hoher Effizienz vollziehen können. Insbesondere wurden

potentielle 53BP1 und RIF1 Interaktoren somatisch untersucht durch CRISPR/Cas9, da diese beiden Faktoren eine Schlüsselrolle während CSR spielen. In dieser Studie wurde herausgefunden, das ZMYND8 GLT und CSR reguliert, indem es 3'RR am *IgH* bindet und moduliert. Außerdem wurde PDAP1 unabhängig als neuer Faktor, welcher für effiziente CSR benötigt wird durch Modulierung von *Aicda* (AID) mRNA-Leveln.
Zusammenfassend tragen die Ergebnisse dieser Arbeit zum weiteren Verständnis der Regulation von Antikörper-Isotyp-Differenzierung und B-Zellenaktivität während einer Immunantwort bei.

1. INTRODUCTION

1.1 The immune system

Mammals have evolved a highly complex and effective immune system in order to protect their organism against foreign pathogens, allergens and toxins. Specifically, the immune system presents innate and adaptive mechanisms to elicit a response. The innate immune system is the first line of defense against antigens and it is characterized by the presence of germline encoded membrane-bound receptors and cytoplasmic proteins that recognize common microbial patterns, known as pathogen-associated molecular patterns (PAMPs)[1]. On the other hand, the adaptive immune system is constituted by lymphocytes (B and T cells) bearing antigen specific receptors generated by somatic recombination or gene rearrangement[2]. These antigen specific cells can persist in the body in order to mount a rapid and robust immune response upon future infection with the same pathogen, a process called immunological memory[3]. Importantly, the two systems are coordinated: while the innate system is rapidly activated upon pathogen encountering, the adaptive system is effective at a later stage because it requires time to generate and expand antigen specific cells that can effectively neutralize it. Moreover, components of the innate response contribute to the activation of lymphocytes, which can in turn make use of the innate mechanisms to mediate their effector function. Importantly, the innate and adaptive systems are tightly regulated in order not to recognize self-antigens that would lead to autoimmune diseases upon degradation[4]. Therefore, the function, modulation and synchronization of adaptive and innate mechanisms are crucial for the establishment of effective immune responses.

1.2 B-lymphocytes and antibody structure

B cells are a central component of the Adaptive Immune System due to their potential to produce antigen-specific antibodies with effector function. Immunoglobulins (Igs) are roughly Y-shaped molecules composed of two identically heavy (IgH) and two identical light (IgL) polypeptide chains bound by disulfide bridges. Each of these chains present a variable region that is responsible for recognizing and binding the antigen, and a constant part (**Figure 1.1**)[5,6]. The variable region is codified by variable (V), diversity (D) and joining (J) genes that get rearranged through a process called V(D)J recombination during B cell development to generate functional antibodies with different antigen specificities. This process is mediated by recombination activating genes-1 and -2 (RAG1 and RAG2) that recognize specific sequences called recombination signals (RSS). Rearrangements always occur in between two gene segments flanked by RSS[7].

In humans, V(D)J recombination has been reported to produce at least 10^{16} antibody variants[8]. Igs further diversify over the course of an immune response due to physiological point mutations of the variable regions by somatic hypermutation (SHM), leading to the full antibody repertoire. On the other hand, the constant part of the heavy chain (C_H) determines isotype of the Ig, conferring diverse effector properties like activation of the complement (group of plasma proteins that are part of the innate system and participate in the immune response) or binding to Fc receptors[5,6]. Initial studies of the antibody structure described different fragments that were observed upon IgG1 digestion with the enzyme papain: Fab or antigen binding fragment, comprising IgL and the variable region plus the first constant domain of IgH, and Fc or crystallizable fragment that includes the rest of C_H and can be easily crystalized (**Figure 1.1**)[9].

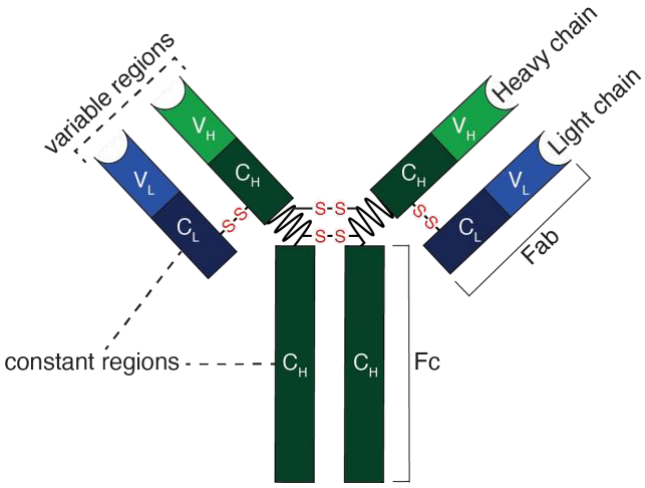

Figure 1.1. Antibody structure.
Schematic representation of the antibody structure as a monomer. The heavy chain (shown in green) and the light chain (colored in blue) are bound by disulfide bridges, indicated in red (S). Both chains exhibit a variable region (V_L and V_H) that determines the antigen specificity and a constant (C_L and C_H) part, which confers effector functions. Fab (antigen binding fragment) and Fc (crystallizable fragment) are indicated.

Importantly, humans and mice present five isotype classes: IgM, IgD, IgG, IgA and IgE. Moreover, IgG is divided in four subclasses that are different in both species, as human IgGs are IgG1, IgG2, IgG3 and IgG4 and murine IgGs are IgG1, IgG2a, IgG2b and IgG3. IgA can be found as IgA1 or IgA2 in humans. The vast majority of B-lymphocytes initially co-express IgM and IgD in their plasma membrane but antigen exposure leads to antibody isotype differentiation to IgG, IgA or IgE by a process called class switch recombination (CSR)[10,11]. At the molecular level, this change occurs at the constant region of the IgH and, therefore, antigen specificity remains the same upon isotype switching. IgM is the first Ig expressed as a monomer on the surface of naïve/unstimulated B cells and it is secreted as a pentamer during an immune response to opsonize the antigen and to activate the complement. IgD is also present in naïve B-

lymphocytes and its effector function is mostly unknown. Monomeric IgG is the most abundant isotype found in the body, constituting 75% of the serum Igs. Although the different subclasses exhibit distinctive functional affinities and exist in different ratios, they are all important to fix the complement and to neutralize virus and toxins. IgA is majorly found in mucosal areas of the body as a dimer, where it prevents bacterial and viral infection. Lastly, IgE is the least abundant Ig but has a strong effector function in allergic reactions and during worm parasite invasions[5,11].

Conclusively, before antigen encountering IgM/IgD B-lymphocytes present a pre-immune repertoire generated by V(D)J recombination. During an immune response, the secondary diversification mechanisms SHM and CSR lead to the formation of antibodies with increased antigen affinity and different effector functions, respectively.

1.2.1 Murine immunoglobulin heavy chain

The IgH is codified by the Immunoglobulin heavy chain locus, *Igh*, that expands over 28 Mb within the chromosome 12 in mice. Mature B cells express only one of the two *Igh* alleles on their surface, by a process known as allelic exclusion that prevents from showing two heavy chains with different antigen specificity within the same cell[12]. Moreover, B-lymphocytes exhibit a variable region resulting from the rearrangement of one V_H, one D_H and one J_H gene segments. Mice have 195 V_H gene segments classified in 16 families based on their sequence homology. Downstream V_H, twelve D_H and four J_H gene segments can be found[13,14]. Specifically, one D_H and one J_H gene segments are rearranged first during VDJ recombination in both *Igh* alleles. Next, a V_H gene segment is joined to the recombined D_H-J_H sequence in one of the alleles, exclusively. If the newly formed V_H-D_H-J_H exon is non-functional the second allele gets recombined[15]. Following the variable region in the *Igh* locus, there is an enhancer element called Eμ that is crucial for supporting efficient *Igh* expression, VDJ recombination and early B cell development[16–19]. Then, the constant part (C_H) follows, consisting of seven different transcription units conformed by a cytokine-inducible promoter, an intervening exon (I-exon), an highly repetitive intronic switch (S) region and a constant gene segment (C) (**Figure 1.2**)[20]. Importantly, IgD results from alternative splicing of a long primary transcript containing both Cμ and Cδ genes[21]. Additionally, the 3' regulatory region (3'RR) is a super-enhancer localized downstream Cα, the last constant gene. This element is conformed of four DNase I hypersensitive sites (hs) and is required for late B cell differentiation, CSR and SHM[22–27] (*see section 1.6*). The 3'RR is followed by additional hs regions (hs5-8) (**Figure 1.2**). Lastly, a putative enhancer element, Eγ, has been identified in between Cγ1 and Sγ2b in developing B cells[28–30]. Although its function

in mature B cells has not been characterized yet, recent literature showed that this enhancer is required for efficient *Igh* transcription and CSR to some isotypes[31].

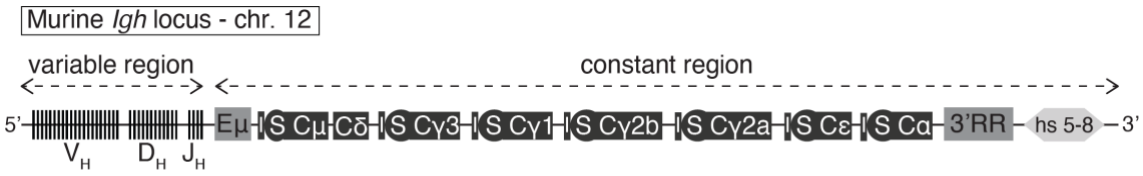

Figure 1.2. The murine immunoglobulin heavy chain locus (*Igh*).
Scheme representing the structure of the murine *Igh* locus (not in scale). The variable region comprises V_H, D_H and J_H gene segments. The constant region has independent transcription units (shown in black) that consist on an intervening exon (I), a highly repetitive element (switch region, S) and a constant gene (Cμ, Cδ, Cγ3, Cγ1, Cγ2b, Cγ2a, Cε, Cα). The *Igh* locus exhibits two enhancer elements, Eμ and the 3' Regulatory Region (3'RR), shown in dark grey. The 3'RR is followed by additional DNase I hypersensitivity sites (hs 5-8).

1.2.2 Murine immunoglobulin light chain

Whereas IgH is codified by a single locus, IgL can result from two independent loci, the kappa (*Igκ*) or the lambda (*Igλ*) locus. As in the case of the *Igh* locus, allelic exclusion leads to the expression of only one allele of either *Igκ* or *Igλ* on the surface of mature B-lymphocytes[15]. Of note, human and murine B cells greatly differ in their expression of *Igκ* versus *Igλ*, as humans present a ratio of 1.5:1 of *Igκ* to *Igλ* and mice have a ratio of 19:1[32].

The murine *Igκ* locus (chromosome 6) has a variable region with more than 100 $V_κ$ segments that, as in the case of V_H, can be classified in different families. Directly downstream, there are five $J_κ$ regions followed by one $C_κ$ that is located in between two enhancer regions: $Ei_κ$ (intronic kappa enhancer) and $3'E_κ$ (3' kappa enhancer) (**Figure 1.3**). Importantly, several of the $V_κ$ murine genes can be found in opposite transcriptional orientation to their associated $J_κ$-$C_κ$ segments[33]. As a consequence of this orientational structure of the *κ* locus, a primary rearrangement of one $V_κ$ with a $J_κ$ does not result in the deletion of the gene segments that are in between. If this first recombination is unproductive or results in a self-reactive B cell receptor (BCR), further rearrangements can take place, since there are additional $V_κ$ segments available. When all recombination possibilities are exhausted, the other *κ* locus in the second chromosome gets rearranged and, upon failure, the *λ* chain gets ultimately recombined. Moreover, another element in the *κ* locus, the recombining segment (RS), plays an important role in this sequential rearrangement. Specifically, it recombines with sequences bearing a specific motif that is present within the $J_κ$-$C_κ$ intron and in the RSS (recombination signal sequences

recognized by RAG) located at the 3' of V_κ. As a consequence of this recombination, the first *κ* locus is discarded and rearrangement starts in the other locus[15,33].

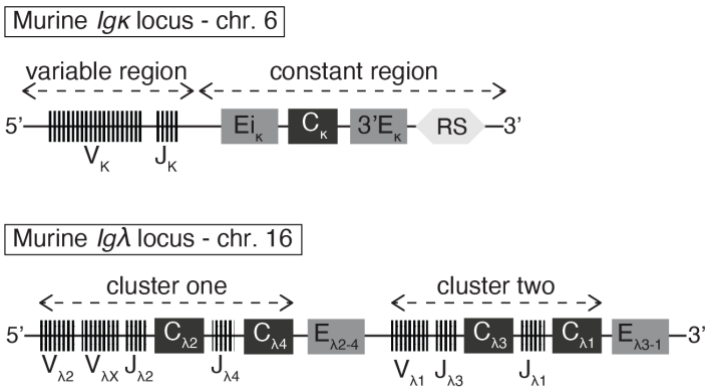

Figure 1.3. The murine immunoglobulin light chain loci (*Igκ* and *Igλ*).
Schematic diagram of the murine *Igκ* (up) and *Igλ* (down) loci. The *Igκ* locus has a variable region followed by a constant part that includes one constant gene segment (C_κ, black) surrounded by two enhancer regions, Ei_κ and $3'E_\kappa$ (dark grey). RS (recombination segment). The *Igλ* locus is configured in two V_λ-J_λ-C_λ clusters. Constant gene segments are depicted in black. There are two enhancer elements, located downstream each cluster (E_λ2-4 and E_λ3-1, dark grey).

In mice, the *Igλ* locus is found in chromosome 16 and in the case of C57BL/6 it is organized in two V_λ-J_λ-C_λ clusters that present the same transcription orientation. Specifically, two V_λ gene segments associate with two functional J_λ-C_λ gene pairs and the third V_λ gene is joined to a second J_λ-C_λ cluster consisting of two J_λ and two C_λ (**Figure 1.3**). Rearrangement happens within the clusters and not in between them and each cluster is flanked by a downstream enhancer (E_λ2-4 and E_λ3-1, lambda 2-4 and lambda 3-1 enhancers, respectively)[33].

1.3 Murine B cell development

B-lymphocytes derive from hematopoietic stem cells (HSCs) that are self-renewing and multipotent cells found in the fetal liver and in the bone marrow (BM) after birth. There are two distinct B cell populations: B-1 and conventional B-2 lymphocytes. B-1 cells mainly originate from the fetal liver and are involved in protecting the organism against self-antigens. On the other hand, B-2 cells principally derive from the BM and participate in the adaptive immune response[34]. In the BM, these B-2 lymphocytes can be induced to differentiate into multipotent progenitor cells (MMPs), which lack the self-renewing capability but can still give rise to all lineages of a tissue or organ. Next, MMPs generate common lymphoid (CLPs) and common myeloid (CMPs) progenitor cells. CMPs are the precursors of erythrocytes, megakaryocytes, granulocytes and macrophages and CLPs lead to B and T lymphocytes and natural killer (NK) cells[35]. Specifically, during B cell

maturation, CLPs differentiate into pre-progenitor B cells (pre-pro-B cells), which are characterized by the expression of the markers CD45R/B220 and CD43 on their surface and have germline Ig genes. Next, progenitor B cells (pro-B) cells express the marker CD19 together with CD45R/B220 and present rearranged D_H-J_H gene segments in their *Igh* loci that express a μ chain. Pro-B cells also express immunoglobulin-α, -β and calnexin proteins, which are signaling components that associate with the BCR in mature B cells. At the precursor stage, pre-B cells become $CD43^-$ and finalize V_H-D_H-J_H recombination in one of the *Igh* alleles, which associates with the so called "surrogate light chain" (SLC)[36–38]. Therefore, expression of Igμ and its binding to the SLC, Igα and Igβ lead to the generation of the pre-BCR (pre-B cell receptor) that is presented on the surface of pre-B cells. If the newly recombined IgH chain successfully pairs with the SLC, pre-B cells continue to rearrange the V and J gene segments of their IgL chains, leading to the replacement of the SLC with an Igκ or Igλ chain and assembly of the pre-BCR. When this pre-BCR results non-functional or autoreactive, further recombination events can occur to substitute it and this process is known as receptor editing. Upon failure to rearrange and express a productive pre-BCR, autoreacting pre-B cells undergo apoptosis and are negatively selected. On the other hand, when recombination and pairing of *Igh* and *Igl* chains are successful, an IgM antibody is expressed on the surface of B cells that are known as immature B-lymphocytes[37,38]. Importantly, the newly formed non-autoreactive BCR expressed in immature B cells needs to achieve a signaling threshold in order for differentiation to continue. This process is known as positive selection, determines the survival of B-lymphocytes in the periphery and it is mainly mediated by the binding of BAFF (B-cell activator of the TNF-α family) to its receptor (BAFF-R)[12,15,36]. Immature B-lymphocytes exit the BM and migrate to the spleen, where they develop in a stepwise manner. Consequently, 3 independent populations known as transitional 1, 2 and 3 (T1, T2 and T3) can be distinguished based on their surface expression of CD23 and IgM: T1 are IgM^{high} and $CD23^-$, T2 are IgM^{high} and $CD23^+$ and T3 are IgM^{low} and $CD23^+$. Importantly, IgD is co-expressed with IgM during late transitional stages as a result of alternative splicing of the *Igh* transcript. Immature B cells express high levels of IgM, while mature naïve B cells present increased IgD and reduced IgM levels[39]. In the spleen, immature transitional B-lymphocytes differentiate into naïve follicular (FO) or marginal zone (MZ) B cells which can be phenotypically distinguished by their surface expression of IgD, CD21 and CD23: MZ are IgD^{low} $CD21^+$ $CD23^-$ and FO are IgD^{high} $CD21^-$ $CD23^+$. Splenic MZ B-lymphocytes rapidly participate in the immune response upon infection independently of their interaction with T cells by secreting IgM molecules and ultimately differentiating into short-lived plasma cells (SLPCs)[40–42]. Within the follicles of secondary lymphoid organs, mature naïve B cells can

mediate a T cell dependent response to pathogenic infection. Specifically, they can be found in anatomically distinguishable functional compartments known as germinal centers (GCs) that are formed roughly 1 week after antigen encountering[43]. There, B cells undergo SHM to further diversify the B cell repertoire, cell proliferation, and affinity maturation, which is a process that mediates the selection of clones that present the highest affinity for the antigen. Moreover, GC is the place where CSR takes place and GC murine B cells are phenotypically characterized by the expression of CD95/Fas and the downregulation of CD38. Highly antigen-specific class-switched B-lymphocytes exit the GC as plasmablasts (PB) that are the precursors of antibody-secreting plasma cells. Importantly, commitment to the plasma cell lineage takes place already in the GC, when B cells express BLIMP-1 (B-lymphocyte-induced maturation protein-1, codified by *Prdm1* gene) that acts as a transcriptional repressor controlling B cell identity. Memory B-lymphocytes also result from a GC reaction and will rapidly differentiate into long-lived plasma cells (LLPCs) or re-enter the GC for further diversification upon re-encountering the same antigen. Finally, these non-dividing long-lived plasma cells migrate back to the BM, where they locate in specialized niches (**Figure 1.4**)[40,43,44].

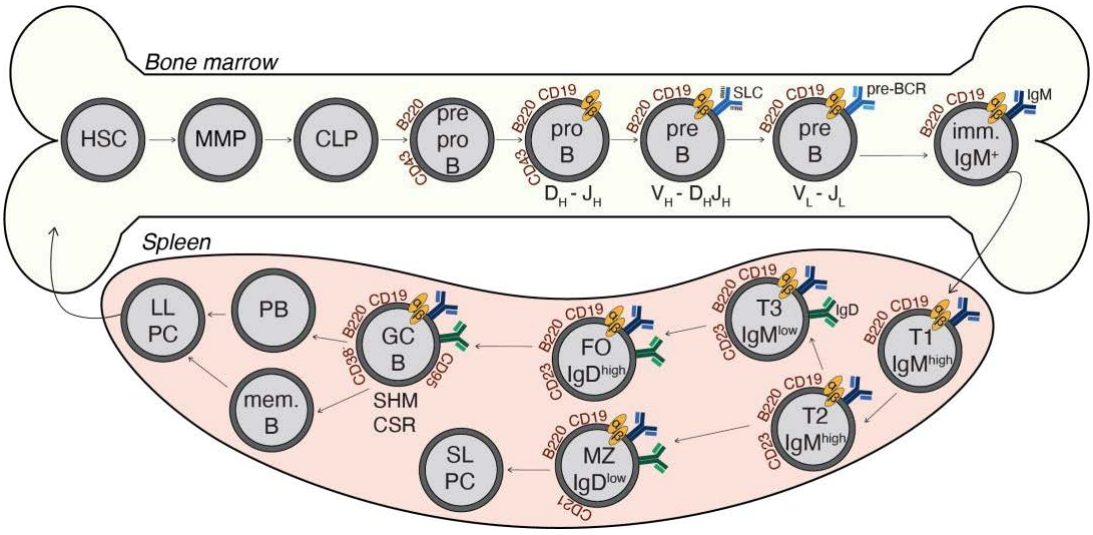

Figure 1.4. Murine B cell development.
Diagram depicting the different stages of B cell maturation in the bone marrow (up) and in the spleen (bottom). HSC, hematopoietic stem cell; MMP, multipotent progenitor; CLP, common lymphoid progenitor; pre-pro-B, pre-progenitor B cell; pro-B, progenitor B cell; pre-B, precursor B cell; SLC, surrogate light chain; pre-BCR, pre-B cell receptor imm. IgM$^+$, immature IgM positive; T1, T2, T3, transitional B cell populations; FO, follicular; MZ, marginal zone; SLPC, short-lived plasma cell; GC B, germinal center B cell; SHM, somatic hypermutation; CSR, class switch recombination; mem. B, memory B cell; PB, plasmablast; LLPC, long-lived plasma cell. IgM and IgD antibody molecules are shown in blue and green, respectively. Surface markers are indicated in red.

1.4 Somatic hypermutation (SHM)

In order to increase antigen specificity, GC B cells undergo SHM, a process that generates single base mutations as well as stochastic base insertions or deletions in the variable regions of IgH and IgL chains. SHM occurs at a frequency of 10^{-5} to 10^{-3} per bp per generation in mice and can be detected from 150-200 bp downstream the IgV promoter up to 1.5-2 kb in the rearranged variable region, including the 3' intron before the C part[45]. Mutations do not happen in a random manner, as there is a preferential "hotspot" or conserved sequence RGYW/WRCY (W = A/T, R = A/G, Y = C/T) that is recognized by the enzyme AID (activation-induced deaminase)[46–49]. AID is exclusively expressed in activated B-lymphocytes and deaminates cytosine (C) to uracil (U) bases in single-stranded DNA (ssDNA)[47,50–52] (**Figure 1.5**). Although AID functions during CSR and SHM[53], the outcomes following U incorporation are different in both molecular mechanisms.

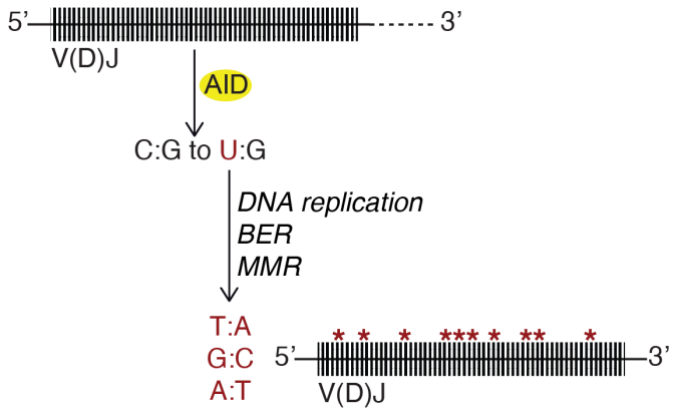

Figure 1.5. Somatic hypermutation.
SHM occurs within the V(D)J region of the heavy and light chains. AID (shown in yellow) enzymatic activity leads to the deamination of cytosine (C) to uracil (U). U:G mismatches are then processed by DNA replication or base excision repair (BER) and mismatch repair (MMR) pathways, leading to transition (C:G to T:A) and transversion (C:G to G:C or A:T) mutations.

Specifically, during SHM there are three possible processes that can happen. First, U may remain in the DNA and act as a template for the incorporation of adenine (A) in the newly synthetized strand during DNA replication, leading to a transition mutation from C:G (guanine) to T (timine):A. Alternatively, U can be recognized by UNG (uracil DNA-glycosilase) or by MSH2/MSH6 (homologue of bacterial MutS 2/6) proteins[54–57]. UNG activity catalyzes the removal of U, which generates an abasic site that is recognized by APE2 (apurinic/apyrimidinic endonuclease 2) and subsequently repaired by the BER (base excision repair) pathway. However, mutagenic repair often happens following UNG activity in AID target spots where the removal of U leads to a non-template site that

can be filled in by polymerases. During SHM, REV1 and polη mediate DNA synthesis and catalyze transversion (from C:G to G:C and A:T) or transition mutations (from C:G to T:A), respectively. Of note, most A:T mutations happening in SHM are dependent on MSH2/MSH6 and EXO1 (exonuclease 1) proteins, which are part of the MMR (mismatch repair) pathway. As in the case of BER, a non-canonical version of this MMR pathway functions upon AID targeting, and EXO1 activity resects the DNA generating a gap, which is filled by polη and other low-fidelity polymerases that ultimately introduce mutations[54–57]. The reason why mutagenic versions of these physiological DNA repair mechanisms function during SHM is yet to be discovered. Moreover, the factors and molecular processes regulating AID targeting are still being investigated.

1.5 Class switch recombination (CSR)

CSR is a deletional somatic recombination process that replaces the constant μ region (Cμ), which is expressed in mature naïve B cells, with one of the downstream constant genes, thus altering the antibody molecule's effector function without affecting its specificity. At the molecular level, CSR occurs between internally repetitive DNA elements that precede each constant gene, the so-called switch (S) regions. CSR is initiated by cytokine stimulation that promotes cell proliferation and germline transcription (GLT) across the switch regions and constant genes. This process creates stretches of ssDNA that are ideal substrates for the B cell specific enzyme AID. As in SHM, AID activity generates U:G mismatches. However, during CSR, these mismatches eventually lead to double-strand breaks (DSBs) formation within the switch regions during the G_1 phase of the cell cycle. DNA damage response (DDR) factors are then recruited to mediate the productive ligation of DSBs in Sμ and one of the downstream S regions by classical non-homologous end-joining (c-NHEJ), a DSB repair process that directly ligates broken ends after minimal DNA end processing (**Figure 1.6**). Therefore, CSR is a multistep process that involves the formation and repair of DSBs at the *Igh* locus. These steps are explained in detail below.

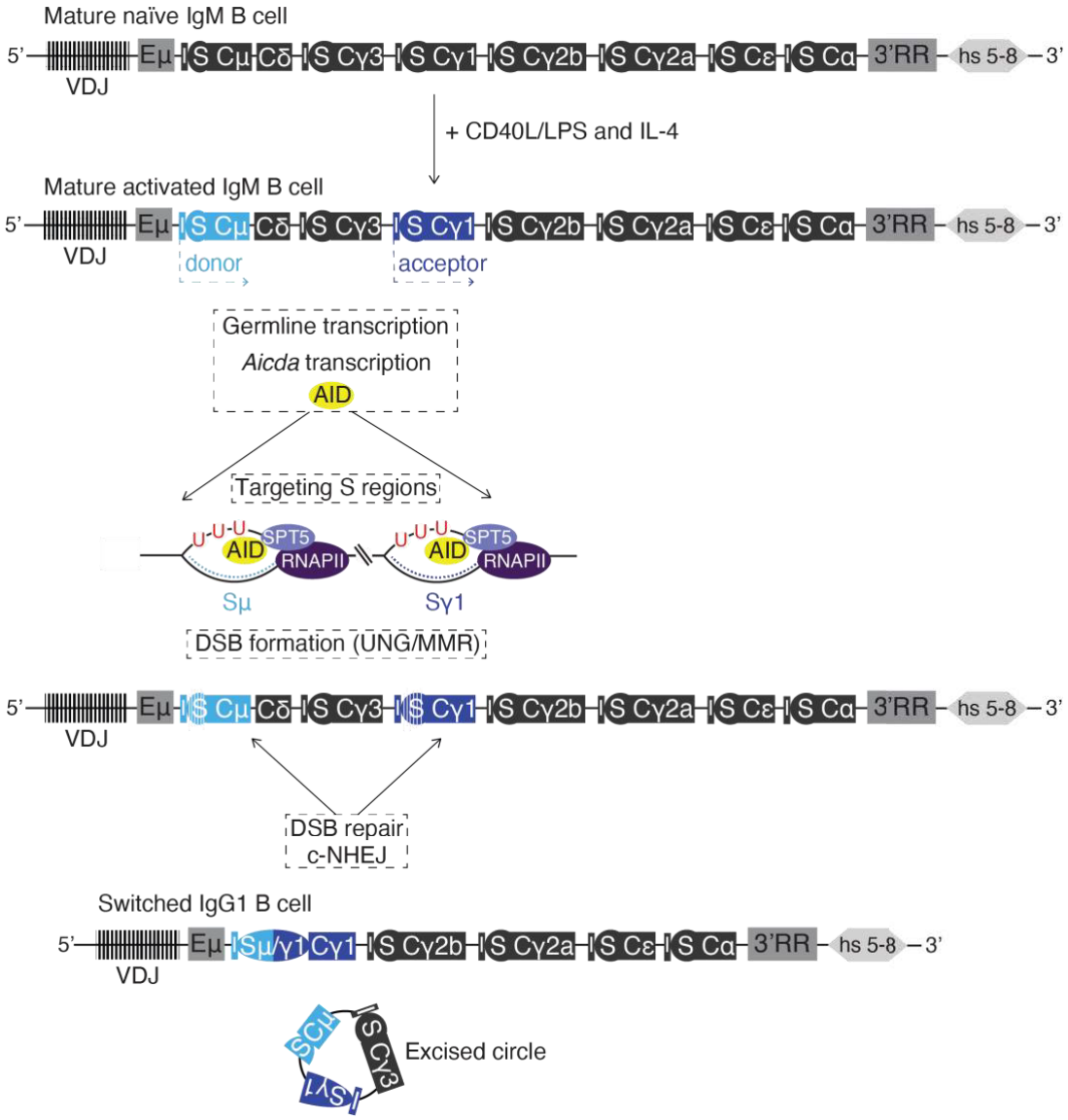

Figure 1.6. Class switch recombination steps.
Diagram of antibody isotype differentiation to IgG1. anti-CD40 or LPS and IL-4 cytokine stimulation induces germline transcription across donor (Sμ) and acceptor (Sγ1) switch regions. As a result, *Aicda* gene gets transcribed and AID is recruited to the S regions through RNA polymerase II and SPT5, generating U:G mismatches that are eventually converted into DSBs. DSBs in the donor and acceptor S regions are repaired by classical non-homologous end-joining (c-NHEJ), leading to CSR to IgG1. Intervening sequence is looped out as a circular extrachromosomal DNA (excised circle).

1.5.1 Germline transcription (GLT)

GLT defines the process of transcription across the individual units that are present in the C part of the *Igh* locus, from the cytokine-inducible promoter located 5' of the I-exon, through the repetitive S region until the 3' of the constant gene. As a result, unprocessed germline transcripts (GLTs) are generated, which are then spliced and polyadenylated, leading to processed/mature products. Transcription happens constitutively across Sμ, also known as the "donor" region and it is induced in a cytokine-dependent manner in

the “acceptor” elements (Sγ1, Sγ3, Sγ2a, Sγ2b, Sε, Sα). Importantly, GLTs do not code for any protein, but their induction in the acceptor region is directly related to antibody differentiation to specific isotypes[20,58]. *In vivo*, CSR can occur in a T cell-dependent (TD) or -independent manner (TI). TD CSR is mediated by the interaction of CD40, which is present on the surface of B-lymphocytes, with CD40 ligand (CD40L) that is expressed by T cells. On the other hand, TI CSR is induced by bacterial or viral antigens that crosslink the BCR and stimulate toll-like receptors (TLRs) in B cells. Both TD and TI responses can be mimicked *in vitro* by treatment with anti-CD40 antibodies or with bacterial lipopolysaccharide (LPS), respectively. Specifically, LPS can bind to TLR4 through its lipid A moiety and to the BCR through its polysaccharidic domain. Furthermore, *in vitro* B cell culturing with different cytokines’ combinations recruit transcription factors and induce GLT in specific acceptor regions that result in CSR to particular isotypes, mimicking TI and TD responses[59]. Consequently, LPS leads to Sγ2b and Sγ3 transcription and subsequent CSR to IgG2b and IgG3, respectively. LPS or anti-CD40 and IL-4 (interleukin-4) stimulate Sγ1 and Sε GLT and expression of IgG1 and IgE on the surface of B-lymphocytes[60–63] TGFβ (transforming growth factor β) activates downstream target proteins that bind the promoter of Sα and Sγ2b, regulating antibody isotype differentiation to IgA and IgG2b[64–66]. IL-5 (interleukin-5) has been shown to increase IgA switching together with LPS and TGFβ[67,68]. Importantly, IL-4 and IL-5 are cytokines secreted by T cells that activate tyrosine kinase cascades, controlling gene expression, viability and proliferation. Similarly, IFN-γ (interferon-γ) is produced by T cells and binds to its receptor to induce signaling and transcription, stimulating the production of Sγ2a and IgG2a isotype differentiation[69]. Lastly, BAFF (B cell activating factor of the TNF-f family) supports proliferation and intracellular signaling and has been reported to enhance CSR to IgA and IgG2b[67].

1.5.2 AID function in CSR

Importantly, cytokines are not only required to induce GLT during CSR, but also *Aicda* gene transcription that peaks at around 48 h post-activation (*see section 1.8.2*)[10]. Although S regions differ in sequence in length (1-12 kb), they are G-rich and contain recurring DNA motifs including hotspots for AID targeting, leading to multiple deaminated U bases in both DNA strands[57,70,71]. These hotspots are enriched in the central part of S regions and become infrequent in the 5’ and 3’ extremes. As in SHM, U residues are removed from the DNA by UNG, generating an abasic site. During CSR, this abasic site is recognized by the apurinic/apyrimidinic endonuclease-1 APE1, which generates a nick or single strand break (SSB). The formation of a proximal additional nick on the opposite strand leads to a DSB. Since AID hotspots are less frequent outside the core of S

regions, there is a decreased probability of forming DBSs and mainly SSBs are present. In this case, MSH2/MSH6 bind U:G mismatches that are neighboring those SSBs and recruit the exonuclease EXO1 that excises the DNA from the closest 5' SSB to the mismatch, leading to the formation of a DSB with a 5' overhang. Importantly, most of the DSBs created within the S regions contain overhangs, which should be processed into blunt ends in order to allow for efficient ligation of two independent regions. This can be done by exo- or endo-nucleases that resect ssDNA and the subsequent activity of polymerases that fill-in the gap[10,20,71].

1.5.3 AID targeting to S regions

Although the precise molecular mechanism by which AID access the S regions is still not fully elucidated, there are several hypotheses described in literature and there is a clear link in between GLT and AID targeting. First, transcription across the S regions exposes stretches of ssDNA, which are the substrates of AID. Moreover, GLTs are G-rich and can form a R-loop (RNA:DNA hybrid) with the non-template strand. As a consequence, the top/template strand of the DNA becomes accessible for AID targeting. Importantly, AID has been shown to equally deaminate both strands *in vivo* but transcription across the S regions has only been observed in one direction, creating controversy in these models[70,72]. However, recent data uncovered a novel role for the core RNA exosome complex, which associates with AID during CSR and degrades the transcript RNA associated to the template strand, resulting in the exposure of stretches of ssDNA[73]. Furthermore, although GLT splicing is essential for efficient CSR to happen, the reason for this is yet to be discovered. Since splicing happens co-transcriptionally, it is possible that the splicing machinery contributes to AID targeting or that it is required for R-loop formation[74]. Moreover, AID has been shown to bind G-quadruplex-like RNA structures, formation of which depends on the nucleotide sequence. Interestingly, the spliced-out intronic S region of GLTs that is G-rich can fold into a G-quadruplex configuration, which interact with soluble AID. AID-RNA complexes can bind the S regions because of the complementarity of the G-quadruplex and its template strand. Alternatively, and considering the existence of R-loops due to GLT, it is also possible that transcripts are processed and released but part of the spliced-out intronic S region remains conforming the DNA:RNA hybrid. These parts that do not anneal with the DNA could then form G-quadruplex structures that recruit AID[58,75].

Additionally, RNA polymerase II (RNAPII) is stalled at S regions, possibly due to R-loop formation. Stalled RNA polymerase II (RNAPII) interacts with SPT5 (suppressor of Ty 5 homolog), which also associates with AID in B cells undergoing CSR. Induction of Us

resulting from AID function are enriched within the 5' end of S regions, leading to the conclusion that AID is recruited to RNAPII during transcription initiation or early elongation[76]. Interestingly, AID phosphorylation at serine-38 (S38) supports its interaction with APE1, which is recruited to the DNA and amplifies the density of DSBs[77]. This AID post-translational modification also mediates its association with RPA, which binds to and stabilizes stretches of ssDNA, favoring AID targeting and the recruitment of DSB repair factors[78–80].

An additional requirement for AID targeting is DNA accessibility, which is determined by specific histone modifications at the *Igh* locus. Sμ donor region has a constitutively poised state, exposing H3K4me3 and H3K9ac marks that increase upon B cell activation. Contrary, constant and acceptor S regions are inaccessible, exhibiting repressive chromatin modifications like H3K36me3 and H4K20me1 (constant regions) H3K9me3 and H3K27ac (acceptor S regions)[81–84]. Cytokine stimulation determines which specific acceptor S region will be transcribed, promoting the removal of these repressive marks and inducing H3K9ac, H3K14ac, H3K27ac, H4K8ac and H3K4me3. A major regulator in this process is PTIP (pax interaction with transcription-activation domain protein-1), a component of the mixed-lineage leukemia-like (MLL) complexes[85,86]. Specifically, PTIP recruits histone methyltransferases that catalyze the formation of H3K4me3. Moreover, R-loops generated during CSR accumulate combinatorial H3K9ac and H2S10 phosphorylation (H3K9AcS10ph), which can be recognized by 14-3-3 adaptor proteins. Importantly, 14-3-3 factors show scaffold functions and have been shown to recruit AID to the S regions[87]. AID targeting is additionally supported by the H3K9me3 mark, which is present in naïve B cells and is retained at low levels in Sμ region upon stimulation. Although this histone modification is characteristic of a repressive sate, it tethers the KAP1-HP1γ (KRAB domain-associated protein 1 and heterochromatin protein 1, respectively) complex that binds AID and facilitates its targeting within the donor region[88].

1.5.4 DSB repair during CSR

Mammalian cells can undergo DSB repair through two main pathways: c-NHEJ and homologous recombination (HR). Whereas the former is active throughout the cell cycle and directly joins the broken ends of a DSB, the latter requires the presence of long stretches of sequence homology in the sister chromatid generated upon replication to copy the information[89–91]. AID-induced DSBs in donor and acceptor S regions at the *Igh* locus are majorly repaired by c-NHEJ, leading to the deletion of the intervening sequence as an extrachromosomal circular DNA and the generation of a newly switched heavy

chain (**Figure 1.6**). The different S regions junctions show limited sequence homology among them and isotype differentiation majorly happens during the G1 and early S phases of the cell cycle, when the sister chromatid is not available for HR[92,93]. DSBs present in S regions are rapidly recognized by the MRN complex, which is composed of MRE11 (meiotic recombination 11 homolog A), RAD50 and NBS1 (Nijmegen breakage syndrome 1). Apart from the MRN complex, DSBs are also recognized by the heterodimer Ku70-Ku80, which is a core component of c-NHEJ and forms a ring that encircles the DNA, preventing the access of nucleases and supporting the recruitment of DNA-PK catalytic subunit (DNA-PKcs)[70–72,94,95]. DNA-PKcs phosphorylates and activates the LIG4-XRCC4-XLF complex, which is composed of essential c-NHEJ factors (DNA ligase IV – X-ray repair cross complementing 4 – XRCC4-like factor) for the ligation of DNA ends and S-S synapsis. NSB1 activates the kinase ATM (ataxia telangiectasia mutated), which phosphorylates threonine and serine residues of several factors involved in the DDR and initiates the G_1/S cell cycle checkpoint until breaks are repaired. Specifically, ATM phosphorylates H2AX (H2A histone family member X, a H2A variant) and this modification, known as γH2AX, spreads in the surroundings of a DSB. Furthermore, 53BP1 (p53-binding protein 1) is an ATM downstream target crucial for c-NHEJ repair and CSR that binds γH2AX, forming foci (focal accumulation) upon DSB formation. During c-NHEJ, phosphorylated 53BP1 protects DNA ends from resection by recruiting its downstream target RIF1 (replication timing regulatory factor 1), favoring S-S efficient synapsis (*see section 1.7*)[96–100].

Furthermore, DSBs can also be repaired by an additional mechanism called alternative non-homologous end-joining (alt-NHEJ) that favors the use of microhomologies at S-S junctions. Consequently, the S-S junctions of B-lymphocytes deficient in c-NHEJ components and factors mediating DNA end protection show increased length of microhomologies, mainly due to DNA end resection of DSBs[94,95,98,101,102]. Since S regions contain internally repetitive sequences but different from each other, there is a high possibility of finding stretches of microhomology within the same S region. This results in “intra-switch” recombination and reduced S donor – S acceptor synapsis, leading to decreased efficiency of antibody isotype differentiation[71].

1.5.5 Cell cycle progression

It is widely accepted that antibody isotype differentiation is linked to cell cycle progression and undivided B cells are CSR-deficient[103–106]. Several studies have shown that DSBs preferentially happen in the G1 phase of the cell cycle. This observation is supported by the fact that both, AID and UNG also predominantly function in this step and also that

unresolved DNA breaks would not scape the G1/S cell cycle checkpoint. Furthermore, AID expression has been linked to cell division[107]. However, there is an additional level of complexity, since further investigations revealed that efficient CSR requires S phase entry, and, possibly the initiation of DNA replication[92,93,108–111]. Recently, it has been shown that activation of replication origins support CSR[112], which is completed in early S phase before the sister chromatid is generated and used for HR[92,113].

1.5.6 Chromatin organization

The *Igh* locus presents a three-dimensional spatial organization that is required for efficient antibody isotype differentiation. Specifically, mature resting B-lymphocytes show intra-chromosomal interaction between the 5'Eμ and the 3'RR enhancer elements (separated by 220 kb) that form a chromatin loop. Upon B cell activation, specific GLT promoters are recruited to the Eμ-3'RR contact region in a cytokine dependent manner, resulting in efficient transcription of targeted acceptor S regions[114,115]. Although the full plethora of factors modulating this spatial organization has yet to be discovered, PTIP has been shown to be required for maintaining the interaction in between the 3'RR and acceptor S regions[85,86]. Additionally, the *Igh* chromatin architectural disposition brings Sμ donor (located immediately downstream Eμ) and acceptor regions in close proximity, allowing synapsis[115,116]. In order for productive CSR to happen, this S-S synapsis are generated in a specific orientation that supports deletional donor-acceptor joining[117]. In this context 53BP1 has been described to support *Igh* architectural structure[118,119], Eμ-3'RR looping and productive orientational joining of AID-induced breaks[117].

1.6 *Igh* 3' regulatory region (3'RR)

The murine 3'RR is a super-enhancer of 28 kb composed of 4 DNase I hypersensitive sites (hs) or core enhancers that can be divided in two structurally distinctive modules. First, hs1,2 is located at the center of a palindromic region flanked by two inverted copies of hs3 (hs3a and hs3b). Secondly, hs4 is found downstream hs3b and is referred as the distal core enhancer[120–127]. Importantly, the DNA sequence that joins these enhancers supports the 3'RR function, most likely by contributing to its architectural configuration[128]. Several mouse models with partial or full deletions of the 3'RR have been generated in the past, showing that the activity of this super-enhancer is crucial for efficient B cell development, CSR and SHM (**Figure 1.7**)[22,23,25–27,128–135] . Furthermore, although some factors have been shown to bind the 3'RR, the precise molecular mechanism of how this region is regulated is still unknown[126].

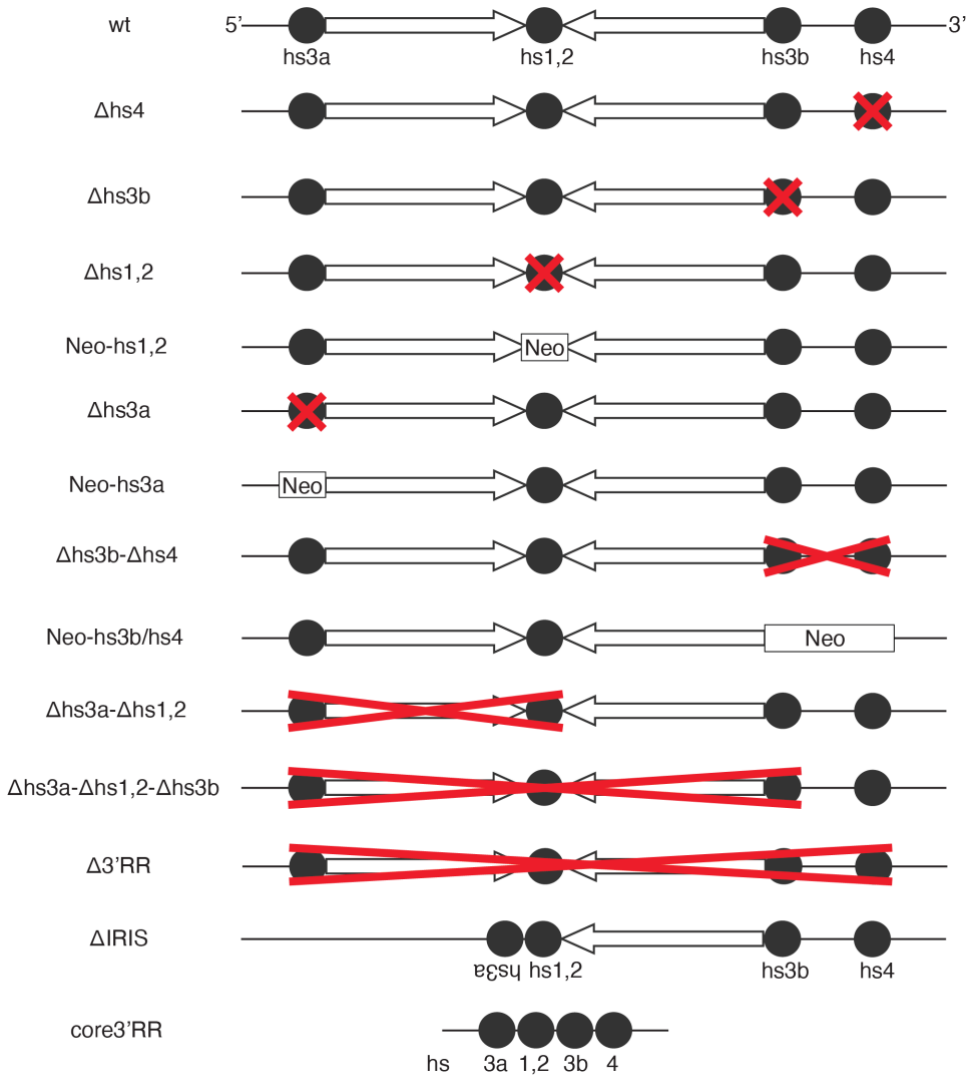

Figure 1.7. The murine *Igh* 3' regulatory region (3'RR).
Scheme representing the 3'RR genomic structure in a wild-type murine B cell (top) and the different mutants that have been described in literature. White arrows depict transcriptional orientation. Deleted regions (Δ) are shown with a red cross. Neo indicates the insertion of a neomycin cassette.

1.6.1 3'RR role in B cell development

The 3'RR is required for late B cell maturation, at the pre-B cell stage and at the transition from immature to mature B cells. Specifically, this super-enhancer supports IgM BCR expression and differentiation towards marginal zone B cells[22–24]. Previously published data indicate that the core enhancer hs4 supports 3'RR activity throughout B cell development. Firstly, hs4 was found to be constitutively active during B cell maturation, while hs3a, hs1,2 and hs3b activity was only detected at late B cell stages[136]. Secondly, single or combined deletion of the core enhancer hs4 together with hs3b (Δhs4, Δhs3b-hs4) lead to decreased number of IgM^+ bone marrow and resting B cells in the spleen, suggesting that this distal module is required for efficient Ig expression[25,129,132]. Finally, Δhs3a-hs1,2-hs3b mice present WT (wild-type) levels of IgM expression in the bone

marrow and spleen as well as unaffected transition of B cells to the marginal and follicular zone[129].

1.6.2 3'RR function during SHM

The 3'RR supports efficient SHM of the *Igh* locus, exclusively. Accordingly, Δ3'RR splenocytes show a nearly complete abrogation in the number of mutations present in the heavy chain, while the light chain remains unaffected[131]. Initial studies proved that deletion of individual 3'RR core enhancers does not affect SHM[26,132,133]. Interestingly, further research clarified that the generation of mutations at the *Igh* locus is dependent on the presence of the proximal 3'RR modules (hs3a, hs1,2 and hs3b), their palindromic structure and also their surrounding sequence of DNA, as core3'RR and ΔIRIS mutants are SHM-deficient[128,129,135].

1.6.3 3'RR function during CSR

The 3'RR regulates CSR during the steps that precede DSB formation at the *Igh* locus, but it is dispensable for repair pathway choice between c-NHEJ and alt-NHEJ and subsequent synapsis of donor-acceptor S regions. Specifically, the 3'RR is required for efficient binding and pausing of RNAPII and transcription of acceptor S regions. Accordingly, 3'RR-deficient cells show decreased R-loop formation and AID targeting upon cytokine stimulation. These cells also display reduced density of H3K4me3 and H3K9ac histone marks, which reflect an open chromatin state[134]. Furthermore, the *Igh* architectural configuration has been shown to be modulated by the 3'RR, as combined deletion of core enhancers hs3b and hs4 impairs Eμ-3'RR loop formation[115].

1.7 53BP1 and RIF1 role in CSR

DSB repair by c-NHEJ is a crucial step for efficient CSR to happen. This pathway relays on the presence of blunt or nearly blunt DNA ends, as limited resection and exposure of microhomology sequences within the S regions skew the repair process towards alt-NHEJ and result in intra-switch recombination[94]. Therefore, factors involved in regulating DNA end processing play a major role in pathway choice between c-NHEJ and alt-NHEJ and productive CSR. 53BP1 is a key modulator of this process since its phosphorylation by ATM upon DNA damage promotes its interaction with RIF1 to prevent DNA end resection by CtIP[96,97,100,137,138]. Recently, two novel complexes, REV7-shieldin (SHLD1, SHLD2, SHLD3) and CST (CTC1-STN1-TEN1), have been identified and suggested to act downstream RIF1 to protect the DNA ends (**Figure 1.8**)[139–146]. Deletion of any of

these factors provides access to proteins mediating DNA end resection and lead to defects in CSR[96,100,137,142,144,146].

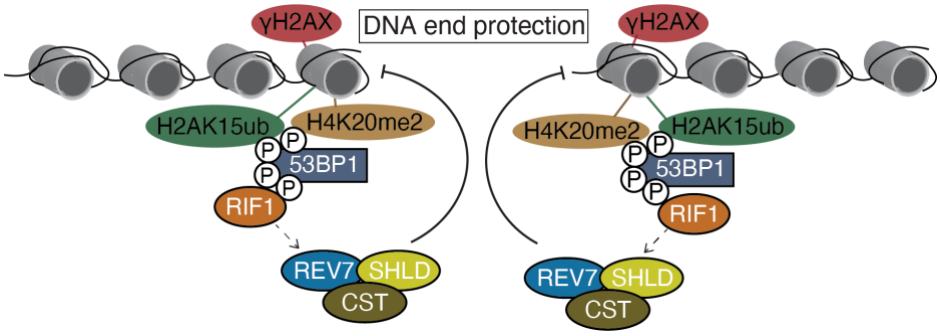

Figure 1.8. DSB DNA end protection.
Diagram representing proteins implicated in DNA end protection. Upon DSB formation, phosphorylated (P) 53BP1 is recruited to damaged sites (marked by γH2AX) through its association to H2AK15ub and H4K20me2. RIF1 interacts with phosphorylated 53BP1 and recruits downstream components (REV7-shieldin and CST complexes) to prevent DNA end resection, ensuring c-NHEJ. Histones are indicated with grey cylinders, that are surrounded by DNA (black line).

1.7.1 53BP1

53BP1 is a large protein of 1972 amino acids. It presents a N-terminal region containing 28 SQ/TQ sites (a serine or a threonine followed by a glutamine residue) that can be phosphorylated by ATM in response to DNA damage[98,99,147,148]. The phosphorylation of these motifs mediates 53BP1 interaction with RIF1 and therefore its function during DNA repair. Accordingly, mutation of these 28 SQ/TQ motifs leads to a decrease in DNA end protection and CSR. Moreover, the N-terminal region of 53BP1 is necessary for chromatin mobility upon DNA damage, although this function is dispensable for antibody isotype differentiation. Next, an oligomerization domain follows, which has been also shown to be required for the focal accumulation of 53BP1 when DSBs are formed and for CSR. Furthermore, 53BP1 is recruited to damaged chromatin via the interaction of its Tudor and UDR (ubiquitination-dependent recruitment) domains with H4K20me2 and H2AK15ub. Both, Tudor and UDR are necessary for DNA damage induced 53BP1 focal accumulation and for preventing DNA end resection of DSBs. During the DDR, ATM also phosphorylates H2A and MDC1 (mediator of DNA damage checkpoint 1) that interact. MDC1 then mediates the recruitment of the ubiquitin ligases RNF8 and RNF168 (ring finger protein 8 and 168, respectively). Lastly, RNF168 catalyzes the formation of the histone mark H2AK15ub and, therefore, supports the association of 53BP1 with the damaged chromatin. The C-terminal part of 53BP1 is characterized by the presence of the BRCT domain, which is dispensable for the focal accumulation of the protein and for CSR[99,147,148].

Apart for supporting DNA end protection, 53BP1 promotes long range S-S synapsis during antibody isotype differentiation. This can be at least partly explained by the facts that 53BP1 is required for efficient chromatin looping of the *Igh* locus and for the correct deletional orientation of S-S junctions during CSR[115,117,118].

1.7.2 RIF1

The mechanism by which 53BP1 mediates DNA end protection upon DSB formation has been studied for many years. Hypothetically, 53BP1 could be preventing the access of nucleases and/or recruiting downstream effectors that actively protect the broken DNA. In this context, RIF1 has been identified as a phospho-dependent 53BP1 interactor that is recruited to DSBs and inhibits DNA end resection. Consequently, RIF1 is required for efficient NHEJ during CSR[96,97,100,137,138]. Furthermore, recent studies have shown that REV7/MAD2L2 (mitotic arrest deficient-like 2) is recruited downstream 53BP1 and RIF1 to counteract DNA end resection upon DSB formation[149,150]. Likewise, the shieldin and CST complexes mediate DNA end protection and are required for efficient CSR. These complexes have been suggested to function downstream 53BP1 and RIF1, but the precise relationship in between all these factors is yet to be defined[139–146]. Although 53BP1 and RIF1 exert common functions in DNA repair, RIF1 is likely to have additional roles because *Rif1* deletion is lethal whereas *Trp53bp1* is not[151]. Indeed, RIF1 has been shown to play a role in transcriptional silencing[152–155], replication origin firing[156,157], repair of stalled forks[151,158], resolution of ultrafine DNA bridges in anaphase[159], and to slow down DNA synthesis upon DNA damage[160]. Murine RIF1 is a protein of 2426 amino acids with a large N-terminal domain containing HEAT repeats that are required for its DSB-dependent interaction with phosphorylated 53BP1 and focal accumulation. The vast majority of the protein is conformed of an intrinsically disordered region followed by a C-terminal domain that is necessary for RIF1 oligomerization and counteract BRCA1 (breast cancer 1; HR factor) activity in the G_1 phase of the cell cycle[137]. Importantly, the contribution of these individual domains to CSR is unknown.

1.8 Physiological relevance of Ig diversification

SHM and CSR are physiological processes necessary for the generation of antigen-specific antibodies with effector properties during an effective adaptive immune response. Consequently, defects in any of these processes can lead to immune defects[161]. Moreover, AID is a B cell specific enzyme induced upon stimulation that is responsible for the generation of mutations and DSBs that are necessary for SHM and CSR to happen, respectively. AID recognizes specific sequence motifs within the

immunoglobulin variable segments and also within the S regions at the *Igh* locus. However, these motifs are also present elsewhere in the genome. Consequently, AID can be targeted to other loci and its mutagenic activity may lead to oncogenic chromosomal translocations or undesired mutations. To prevent this, there are several molecular mechanisms that control AID functionality at the level of transcription, post-transcription, post-translation and subcellular localization. Furthermore, AID-induced DSBs in donor and acceptor S regions are repaired by c-NHEJ during productive CSR events. However, inefficient repair may increase the load of DBSs and alt-NHEJ can make use of sequences microhomologies to fuse DNA ends together. As a result and considering that AID targeting is not exclusive to the S regions, abnormal joining can happen, contributing to translocation events[162–164]. Conclusively, there is a tight modulation at the molecular level of the physiological processes that ensure immunoglobulin diversification while preventing malignant transformation during an adaptive immune response.

1.8.1 Primary antibody deficiencies caused by defects in CSR/SHM

Primary antibody deficiencies (PADs) are considered a group of rare diseases characterized by the inability to support an efficient immune response[165,166]. They are majorly caused by B cell-intrinsic defects, although they can also result from the loss of function of other immune cell types, like T-lymphocytes. Regarding B-lymphocytes, there are syndromes associated to deficiencies in B cell development, survival, proliferation signaling and Ig diversification. Specifically, defects in CSR (often referred as hyper-IgM syndromes) lead to normal or increased IgM levels and reduced IgG, IgA and IgE concentrations in serum. They are usually diagnosed during childhood and they curse with recurrent bacterial and opportunistic infections. Depending on the molecular cause, deficiencies in CSR can be accompanied with defects in SHM. For example, CD40L (expressed in activated T cells) interaction with CD40 (constitutively present in B cells and other immune cell types) is required for GC formation. Mutations in CD40 result in both, CSR and SHM defects and they lead to a rare autosomal-recessive syndrome. Mutations in CD40L (expressed by activated T cells) cause the X-linked form of the disease. Defects in AID abolish CSR and SHM, resulting in autosomal-recessive AID deficiency. Importantly, mutations in the C-terminal region of AID completely abrogate CSR but do not affect SHM. Furthermore, alterations in the nuclear export of AID cause an autosomal-dominant CSR defect, which is clinically less severe than the autosomal-recessive form. Lack of UNG leads to a rare autosomal-recessive CSR deficiency that courses with defects in SHM. Furthermore, inefficient DSB repair can also cause CSR deficiency. For example, functional impairment of the MMR pathway has been linked to

reduced antibody isotype differentiation to IgA and IgG. Likewise, mutations in the DNA damage response and repair factors ATM, RNF168 and MRE11 have been shown impair CSR[161,166,167]. Importantly, there are many patients that present defects in Ig diversification, but the genetic cause is still unknown. Therefore, the identification and molecular characterization of novel players that support CSR and SHM during an adaptive immune response are crucial for the diagnosis and treatment of these patients.

1.8.2 Regulation of AID activity

The *Aicda* gene is encoded in chromosome 6 in mice. It presents four distinctive and conserved DNA regions that are the binding sites for several transcription factors that regulate its expression in B-lymphocytes: Region I functions as a promoter; region II contains silencer elements bound by repressors in memory B cells and non-B cells that restrict AID expression; region III has been shown to support physiological levels of AID expression; region IV and region II present enhancer elements that overcome the silencing of *Aicda* transcription during B cell activation (**Figure 1.9**)[168]. Both, TD and TI signals promote CD40 engagement or dual TLR-BCR crosslinking, leading to activation of the canonical (CD40 and TLR) and non-canonical (CD40 and BCR) NF-κB (nuclear factor kappa beta) signaling cascades. The non-canonical NF-κB p52 subunit binds to the *Aicda* promoter and the canonical NF-κB p65 subunit associates with an upstream enhancer region. LPS binds to TLR4 and the BCR, activating both NF-κB pathways that initiate and sustain *Aicda* expression. Furthermore, IL-4, TGFβ and IFN-γ activate transcription factors that bind and support the induction of *Aicda* gene and that inhibit the activity of silencers restricting *Aicda* expression[10].

Additionally, AID transcription is also dependent on histone modifications of the locus. Specifically, the *Aicda* locus of naïve B-lymphocytes presents low levels of histone H3 acetylation and a hypermethylated promoter. During stimulation, the promoter is demethylated and the locus shows H3K4me3, H3K9ac and H3K14ac modifications. Moreover, *Aicda* transcription elongation is supported by H3K36me3. Therefore, both epigenetic marks and the induction of transcription factors participate in the expression of AID during B cell activation and CSR[168].

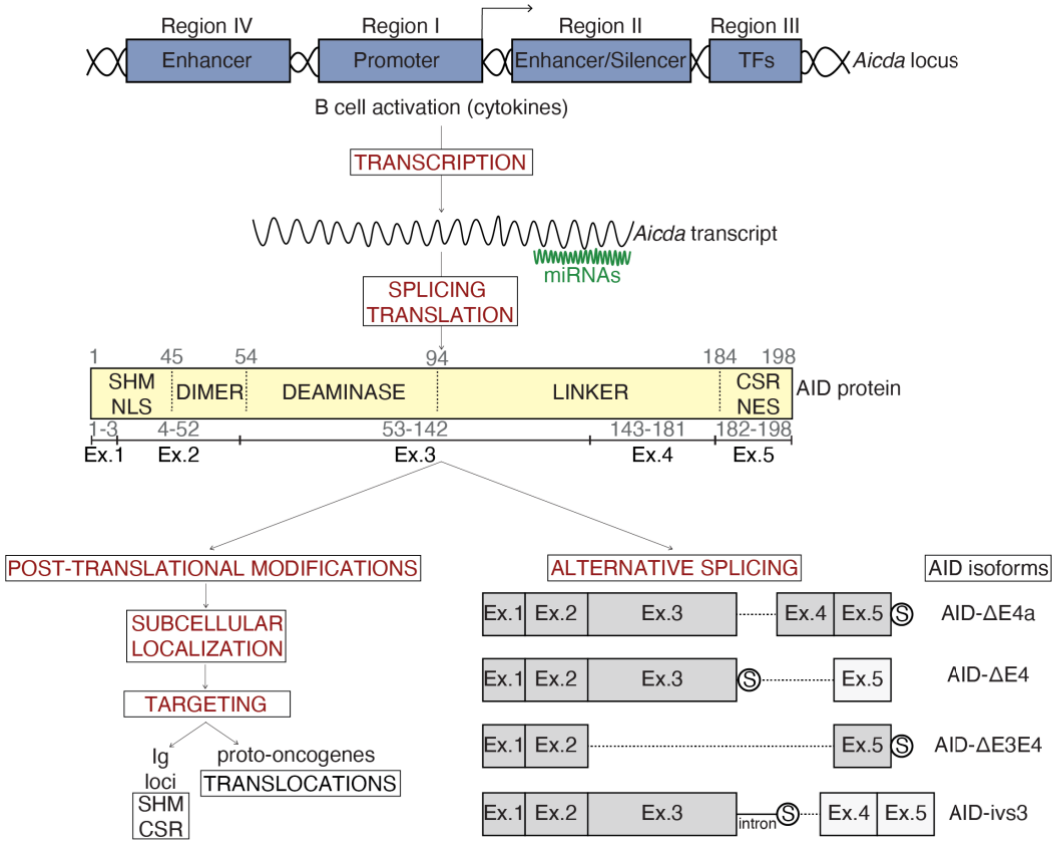

Figure 1.9. Mechanisms of AID regulation.
Diagram depicting the different levels of regulation of AID activity. NLS, nuclear localization signal; NES, nuclear export signal. Representations of the different regions of the Aicda locus, the AID protein domains and the four alternative splicing forms are adapted from previous publications[168,169].

Once transcription is completed, *Aicda* mRNA can be negatively regulated by miRNA species (**Figure 1.9**). miRNAs are small (21-23 nt) non-coding evolutionary conserved RNAs that bind target mRNAs to promote their degradation and inhibit their translation. Specifically, miR-155, miR-181b and miR-361 have been shown associate with the 3'UTR (untranslated region) of *Aicda* transcript to limit its expression[83]. Among all, miR-155 is the most characterized and abundant non-coding RNA that is induced together with AID in activated B cells undergoing CSR. Mutation of the miR-155 binding site in the *Aicda* mRNA leads to an increase in AID protein translation, antibody isotype differentiation and *c-myc/Igh* translocations[170,171]. Therefore, the post-transcriptional regulation of *Aicda* by miRNAs represents a mechanism of control of malignant transformation in B cells.

Aicda transcript can be alternatively spliced, leading to the generation of four different AID isoforms in addition to the full-length (FL) with distinctive exon organization. Although these isoforms have been identified in physiological conditions, they have been shown to be important in patients with B cell malignancies[172,173]. Specifically, AID protein

is organized in 5 functional domains: the N-terminal region codified by exon 1 and part of exon 2 that is required for SHM and contains a nuclear localization signal; a dimerization signal, which corresponds to the rest of exon 2; the functional deaminase activity domain, transcribed from a portion of exon 3; a linker region that is partly codified by exon 3 and also exon 4; and the C-terminal domain, which supports CSR and presents a nuclear export signal (exon 5) (**Figure 1.9**)[168]. The AID-ΔE4a isoform lacks a portion of the linker. The AID-ΔE4 and AID-ivs3 variants present a stop codon after exon 3 and, therefore, the linker and the CSR domain are spliced out. Lastly, AID-ΔE3E4 retains the CSR region but lacks the deaminase domain[169]. Of note, some of these isoforms were initially described to be catalytically active[169], but later reports showed that although *Aicda* undergoes alternative splicing, the resulting isoforms do not code for a functional AID protein[174,175]. Taking all results into consideration, *Aicda* splicing is a critical step for the generation of a functional protein with physiological roles in SHM and CSR.

Furthermore, AID activity is also regulated by post-translational modifications that involve serine and threonine phosphorylation (**Figure 1.9**). Specifically, serine-3 phosphorylation of AID (S3) suppresses its activity *in vivo*. Consequently, S3 mutation results in increased CSR and *c-myc/Igh* translocations, whereas induction of S3 phosphorylation decreases antibody isotype differentiation[176]. On the other hand, phosphorylation of serine-38 (S38) or threonine-140 (T140) positively modulate AID function[80]. S38 is the best characterized AID post-translational modification and it supports AID targeting to the V and S regions during SHM and CSR, respectively. Interestingly, supra-physiological levels of S38 phosphorylation result in increased *c-myc/Igh* translocations but not CSR, suggesting that AID function can be differentially regulated by post-translational modifications[177].

Lastly, AID function is tightly modulated by its subcellular localization (**Figure 1.9**). Indeed, AID is mainly retained in the cytoplasm and its accessibility to the nucleus is restricted in order to minimize undesired off-target deamination. A combination of mechanisms that mediate cytoplasm retention, degradation of nuclear AID as well as active nuclear import and export ensure that only the necessary amount of protein is targeted to the DNA[178–180]

1.8.3 Implications of AID off-targeting and dysregulation

Chromosome translocations result from the joining of DSBs present in non-homologous chromosomes that are in close proximity. These translocations can lead to malignant

transformation when involving proto-oncogenes or tumor suppressor genes. Indeed, many types of B cell lymphomas are caused by translocations that activate proto-oncogenes as a consequence of their binding to the *Igh* locus, which is actively transcribed. For example, the coding region of the proto-oncogene *c-myc* is fused to the *Igh* locus in Burkitt lymphoma, which originates from GC B cells. AID has been shown to target several *non-Ig* genes, including the proto-oncogenes *c-myc*, *Pax5*, *Pim1*, *Bcl6* and *RhoH* and deregulated AID expression causes B cell lymphomagenesis[163,181]. Importantly the repair of AID-induced on- and off-target DSBs breaks also plays a role in chromosomal translocations. The absence of c-NHEJ components can lead to the persistence of unrepaired DSBs that can be processed and aberrantly joined by alt-NHEJ. Accordingly, there is an increase in the level of translocations in Ku-, XRCC4-, LIG4- or DNA-PKcs-deficient B cells. However, despite the higher number in chromosomal translocations, these defects in c-NHEJ factors do not result in tumor onset because cells undergo apoptosis[182–184].

Furthermore, AID activity in chronic myeloid leukemia promotes genomic instability, chemotherapy resistance and progress into blast crisis (incurable and fatal phase of the disease)[185]. Interestingly, AID upregulation has been observed in solid tumors. For instance, AID activity results in point mutations and copy alterations of tumor suppressor genes in gastric cancer[186]. Additionally, *Helicobacter pylori*, which is involved in gastric carcinogenesis, has been shown to induce aberrant AID expression in gastric epithelial cells, resulting in accumulation of undesired mutations[187].

2. GOAL AND SPECIFIC AIMS

CSR is a complex physiological process that involves the formation and repair of DSBs through different molecular mechanisms that have not been fully elucidated yet. Therefore, the aim of this study was to identify novel CSR factors, and characterize their role(s) in the reaction. This was achieved as detailed in the specific aims below:

2.1 Set-up and perform a functional loss-of-CSR screen

To identify novel factors regulating CSR, a functional loss-of-CSR screen by CRISPR/Cas9 somatic targeting was established in the B cell lymphoma line CH12F3, which can undergo CSR to IgA *in vitro*. 53BP1 is a key factor involved in various steps in CSR but the precise molecular mechanisms by which it exerts its functions is still unknown. When this study started, RIF1 had just been identified as a 53BP1 downstream effector in CSR by a mass-spectrometry based approach[96]. Therefore, this list of potential 53BP1 interactors was available and was used to select candidates for the loss-of-CSR assay. Moreover, a similar experimental approach was employed to define the RIF1 interactome in B-lymphocytes activated to undergo CSR, and to test the involvement of the newly identified associating proteins in antibody isotype differentiation.

2.2 Molecular characterization of the newly identified CSR effectors

CSR is a physiological process that is dependent on B cell proliferation and cell cycle progression, efficient AID expression and targeting, DSB formation, chromatin remodeling and DSB repair. Considering that the newly identified effectors could be playing a role in any of these processes, several experimental approaches and functional assays were employed to characterize their functions in CSR.

Therefore, the identification and characterization of novel CSR factors aimed to further define the molecular mechanisms ensuring antibody isotype differentiation and the establishment of an efficient immune response.

3. MATERIALS AND METHODS

3.1 Materials

Mouse lines		
Line	Company/Reference	Cat. No.
C57BL6 (*wt)*	The Jackson Laboratory	664
Rif1$^{FH/FH}$	156	N/A
AID$^{-/-}$	53	N/A
Zmynd8$^{tm1a(EUCOMM)Wtsi}$	European Mouse Mutant Archive	05720
Rosa26^{Flpo}	The Jackson Laboratory	007844
Cd19$^{Cre/+}$	188	N/A
Zmynd8$^{F/F}$	This study	N/A
Zmynd8$^{F/F}$*Cd19*$^{Cre/+}$	This study	N/A
Pdap1$^{F/F}$	This study	N/A
Pdap1$^{F/F}$*Cd19*$^{Cre/+}$	This study	N/A
Cell lines		
Line	Company/Reference	Cat. No.
CH12 (CH12F3) (*wt*, parental)	189	N/A
WTc CH12 (clonal derivative)	This study Daniel B. Rosen	N/A
Aicda$^{-/-}$ CH12	Kerafast	ESP013
Lig4$^{-/-}$ CH12	190	N/A
53bp1$^{-/-}$ CH12	This study Devakumar Sundaravinayagam	N/A
Rif1$^{-/-}$ CH12	This study Matteo Andreani	N/A
Zmynd8-KO$_1$ CH12	This study Daniel B. Rosen	N/A
Zmynd8-KO$_2$ CH12	This study Daniel B. Rosen	N/A
Zmynd8-KO$_3$ CH12	This study Verónica Delgado-Benito	N/A
iMEFs (*wt*, parental)	98	N/A
WTc iMEFs (clonal derivative)	This study Verónica Delgado-Benito	N/A
Random iMEFs (*Rc-1, Rc-2)*	This study Verónica Delgado-Benito	N/A
Zmynd8$^{-/-}$ iMEFs	This study Verónica Delgado-Benito	N/A
Brca1$^{-/-}$ iMEFs	192	N/A
Pdap1$^{-/-}$ CH12 (P1, P2, P3)	This study Verónica Delgado-Benito	N/A

BOSC23 (*wt*)	191	N/A
Antibodies		
Name	Company/Reference	Cat. No.
Donkey anti-goat (HRP)	Gift from Claus Schreidereit's lab	N/A
Goat anti-IgA-PE	Southern Biotech	1040-09
Goat anti-IgG1, Human ads-BIOT	Southern Biotech	1070-08
Goat anti-IgM, Human ads-BIOT	Southern Biotech	1020-08
Goat anti-Kappa-UNLB	Southern Biotech	1050-01
Goat anti-Lamda-UNLB	Southern Biotech	1060-01
Goat anti-mouse Alexa488	ThermoFisher Scientific	A-11029
Goat anti-mouse Lc (HRP)	Jackson ImmunoResearch	115-035-174
Goat anti-rabbit Alexa546	ThermoFisher Scientific	A-11035
Hamster anti-CD95/Fas-FITC Clone Jo2	BD Biosciences	554257
Hamster anti-CD95/Fas-PE Clone Jo2	BD Biosciences	561985
Mouse anti-53BP1 Clone BP18	Millipore/Merck	05-725
Mouse anti-Flag M2	Sigma-Aldrich	F3165
Mouse anti-Flag M2 (HRP conjugated)	Sigma-Aldrich	A8592
Mouse anti-rabbit Lc (HRP)	Jackson ImmunoResearch	211-032-171
Mouse anti-RNA polymerase II CTD repeat YSPTSPS (phospho S5) Clone 4H8	Abcam	ab5408
Mouse anti-βActin Clone AC-15	Sigma-Aldrich	A5441
Mouse anti-γH2AX (Ser139) Clone JBW301	Millipore	05-636
Mouse IgG1-UNLB Clone 15H6	Southern Biotech	0102-01
Mouse IgM-UNLB Clone 11E10	Southern Biotech	0101-01
Rabbit anti-PDAP1	Sigma-Aldrich	HPA050294
Rabbit anti-Rif1	96	N/A
Rabbit anti-ZMYND8	Sigma-Aldrich	HPA020949
Rat anti-AID mAID-2	ThermoFisher Scientific	14-5959-82
Rat anti-B220/CD45R-BV785 Clone RA3-6B2	BioLegend	103245
Rat anti-B220/CD45R-FITC Clone RA3-6B2	BioLegend	103205
Rat anti-B220/CD45R-PE Clone RA3-6B2	BioLegend	103207
Rat anti-CD180 (RP/14)	BD Biosciences	552128
Rat anti-CD19-Alexa700 Clone 6D5	BioLegend	115527
Rat anti-CD19-APC Clone 6D5	BioLegend	115512
Rat anti-CD19-PacBlue Clone 6D5	BioLegend	115523
Rat anti-CD21/CD35-FITC Clone 7G6	BD Biosciences	561769
Rat anti-CD23-PE Clone B3B4	BioLegend	101607
Rat anti-CD3-PE Clone 17A2	BioLegend	100206

Rat anti-CD38-Alexa700 Clone 90	ThermoFisher	56-0381-82
Rat anti-CD40 Clone 1C10	BioLegend	102810
Rat anti-CD43-PE Clone S7	BD Biosciences	561857
Rat anti-IgD-FITC Clone 11-26c.2a	BD Biosciences	562022
Rat anti-IgG1-APC Clone X56	BD Biosciences	550874
Rat anti-IgG2b-PE Clone RMG2b-1	BioLegend	406707
Rat anti-IgG3-BIOT Clone R40-82	BD Biosciences	553401
Rat anti-IgM-FITC Clone II/41	BD Biosciences	553437
Rat anti-IgM-PE Clone R6-60.2	BD Biosciences	553409
Secondary rat	Gift from Fritz Rathjen's lab	N/A
Streptavidin-AP Conjugate	Sigma-Aldrich	11089161001
Streptavidin-APC	BD Biosciences	554067
TruStain fcX Rat anti-CD16/32 Clone 93	BioLegend	101320

Chemical, peptides, recombinant proteins and commercial kits		
Reagent	Company/Reference	Cat. No.
$^{13}C_6$ L-arginine	Cambridge Isotope Laboratories	CLM-2265-H-PK-1
$^{13}C_6$ L-lysine	Cambridge Isotope Laboratories	CLM-2247-H-PK-1
1X PBS	ThermoFisher Scientific	10010031
2-Mercaptoethanol	ThermoFisher Scientific	21985023
3XFlag peptide	Sigma-Aldrich	F4799
4-hydroxytamoxifen (4-HT)	Sigma-Aldrich	H7904
4-nitrophenyl phosphate	Sigma-Aldrich	N2765
Acetic acid	Roth	6755.1
ACK lysis buffer	ThermoFisher Scientific	A1049201
Actinomycin D	Sigma-Aldrich	A9415
Agarose	Roth	2267.4
AllPrep DNA/RNA Mini Kit	Qiagen	80204
Alum $KAI(SO4)_2$	Sigma-Aldrich	31242
Antibiotic Antimycotic	ThermoFisher Scientific	15240062
BAFF (human recombinant)	PeproTech	310-13
Benzonase	Sigma-Aldrich	E1014
Bradford reagent	Bio-Rad	500-0006
BSA	NEB	B900S
BSA (Albumin, IgG-free)	Roth	3737
Carboxyfluorescein succinimidyl ester (CFSE)	ThermoFisher Scientific	65-0850-84
CellTrace Violet (CTV)	ThermoFisher Scientific	C34557
Colcemid	Sigma-Aldrich	10295892001
Complete EDTA-free protease inhibitor cocktail	Roche	11873580001
Crystal Violet	Sigma-Aldrich	C0775
Dialyzed fetal bovine serum	ThermoFisher Scientific	26400-044
Diethanolamine	Roth	0332
DirectPCR Ear Lysis reagent	VWR	402-E

DMEM	ThermoFisher Scientific	41965062
DTT	ThermoFisher Scientific	R0862
Dynabeads M-270 Epoxy	ThermoFisher Scientific	14301
Dynabeads Protein A	ThermoFisher Scientific	10001D
ECL	PerkinElmer	NEL104001EA
EDTA	Roth	8043.2
Ethanol	Roth	P075.1
FBS	Sigma-Aldrich	F7524
FuGENE HD Reagent	Promega	E2312
Gel loading dye (6X)	NEB	B70245
Gentamicin	Lonza	17-519
Gibson Assembly master mix	BioLabs	E2611S
Glutaraldehyde	Sigma-Aldrich	340855
Glycine	Roth	3908
Glycogen	Roth	HP51.3
Gurr Buffer Tablets	ThermoFisher Scientific	10582013
H_3BO_3	Roth	6943
HEPES	ThermoFisher Scientific	15630056
High Sensitivity D1000 Ladder	Agilent	5067-5587
High Sensitivity D1000 Sample Buffer	Agilent	5067-5603
High Sensitivity D1000 ScreenTape	Agilent	5067-5584
HotStarTaq DNA Polymerase	Qiagen	203205
Igepal CA-630	Sigma-Aldrich	I8896
IL-4 (mouse recombinant)	Sigma-Aldrich	I1020
Immobilen PVDF membrane	Millipore	IPVH00010
Isofluran	CP-Pharma	G81G17A
KaryoMAX Giemsa Stain Solution	ThermoFisher Scientific	10092013
KCl	Roth	6781
L-Glutamine	ThermoFisher Scientific	25030024
LPS	Sigma-Aldrich	L2630
Luna Universal qPCR Mastermix	NEB	M3003
Methanol	Chemsolute	143.2511
$MgCl_2$	Sigma-Aldrich	M8266
NaCl	Roth	3957.1
NaOH	Roth	6771
Neon Transfection System, 100µl Kit	ThermoFisher Scientific	MPK10025
non-labeled L-arginine	Sigma-Aldrich	A8094
non-labeled L-lysine	Sigma-Aldrich	L8662
NP-CGG_{16}	Biosearch Technologies	N5055B
NP-IgG1	Gift from Klaus Rajewsky's lab	N/A
NP-IgM	Gift from Klaus Rajewsky's lab	N/A
NP_{22}-BSA	Gift from Klaus Rajewsky's lab	N/A
NucleoSpin DNA Purification Kit	Macherey-Nagel	740499
NucleoSpin Gel and PCR Clean-up kit	Macherey-Nagel	740609

NuPAGE 3-8% Tris-Acetate Protein Gel	ThermoFisher Scientific	EA03752BOX
NuPAGE 4-12% Bis-Tris Protein Gel	ThermoFisher Scientific	NP0322BOX
NuPage LDS Sample buffer	ThermoFisher Scientific	NP0008
NuPage MES SDS Running Buffer	ThermoFisher Scientific	NP0002
NuPage MOPS SDS Running Buffer	ThermoFisher Scientific	NP0001
NuPage Tris-Acetate SDS Running Buffer	ThermoFisher Scientific	LA0041
One Shot Stbl3 chemically competent *E. coli*	ThermoFisher Scientific	C737303
One Shot TOP10 chemically competent *E. coli*	ThermoFisher Scientific	C4040320
Opti-MEM medium	ThermoFisher Scientific	31985062
Paraformaldehyde	Sigma-Aldrich	P6148
PARPi Olaparib/AZD2281, Ku-0059436	Selleckchem	S1060
Penicillin-Streptomycin	ThermoFisher Scientific	15140122
Phenol:Chloroform:Isoamyl alcohol	Roth	A156.3
Phosphatase inhibitor cocktail tablets	Roche	04906837001
Phusion High-Fidelity DNA Polymerase	ThermoFisher Scientific NEB	F530L M0530L
Pierce 16% Formaldehyde Methanol-free	ThermoFisher Scientific	28906
Pierce complete protease inhibitor mini tablets	ThermoFisher Scientific	A32953
Polyvinylpyrrolidone	Sigma-Aldrich	PVP40
Prolong Gold Antifade Mountant	ThermoFisher Scientific	P36930
Proteinase K	Peqlab	3375501
PureLink HiPure Plasmid Midiprep Kit	ThermoFisher Scientific	K210005
Qubit dsDNA HS Assay Kit	ThermoFisher Scientific	Q32851
Qubit RNA HS Assay Kit	ThermoFisher Scientific	Q32852
RapidOut DNA Removal Kit	ThermoFisher Scientific	K2981
Rat anti-CD43 (Ly-48) Microbeads	Miltenyi Biotec	130-049-801
recombinant murine IL-5	Peprotech	215-15
Ribo-Zero Gold rRNA Removal Kit	Illumina	MRZG12324
RIPA buffer	Sigma-Aldrich	R0278
RNeasy Mini Kit	Qiagen	74104
RPMI 1640	ThermoFisher Scientific	21875091
SDS	Roth	EN30.2
SeeBlue Plus2 Pre-stained Protein Standard	ThermoFisher Scientific	LC5925
SILAC-RPMI	ThermoFisher Scientific	88421
Sodium Acetate	Sigma-Aldrich	52889
Sodium Pyruvate	ThermoFisher Scientific	11360039

SuperScript VILO cDNA Synthesis Kit	ThermoFisher Scientific	11754250
TE buffer	ThermoFisher Scientific	12090015
TGFβ-1 (mouse recombinant)	R&D Systems	7666-MB-005/CF
TOPO TA Cloning Kit	ThermoFisher Scientific	450641
Tris	Roth	AE15.3
TRIzol LS Reagent	Invitrogen	10296028
TruSeq Stranded Total RNA Library Prep Kit	Illumina	20020597
TruSeq Stranded Total RNA Library Prep Kit Gold	Ilumina	20020598
Trypan Blue	ThermoFisher Scientific	15250061
Tween	Sigma-Aldrich	P7949
Materials and Equipment		
Name	Company/Reference	Cat. No.
0.45 µM filter	GE Healthcare	10462100
0.50 µM filter	CellTrics	04-0042-2317
0.70 µM cell strainer	Falcon	352350
1 ml syringe	Braun	9161496V
1.5 ml tubes	Sarstedt	72.690.001
10 cm dishes (cell culture)	Sarstedt	83.3902
10 ml pipettes	Sarstedt	86.1254.001
10 ml syringe	Braun	4606198V
12-well plates (cell culture)	Greiner	665180
15 ml tubes	Falcon	352096
25 cm dishes (cell culture)	Sarstedt	83.3903
25 ml pipettes	Sarstedt	86.1685.001
3 ml syringe	Braun	4616025V
4200 ScreenTape System	Agilent	G2991AA
5 cm dishes (cell culture)	Corning	403196
5 ml pipettes	Sarstedt	86.1253.025
50 ml pipettes	Sarstedt	86.1256.001
50 ml syringe	BD Plastipak	300865
50 ml tubes	Falcon	352070
6-well plates (cell culture)	Sarstedt	83.3920
96-well plates (high binding) (ELISA)	Greiner	675061
96-well plates (RT-qPCR)	Applied Biosystems	4346907
96-well plates F bottom (cell culture)	Greiner	655083
96-well plates lids (qRT-PCR)	Applied Biosystems	4311971
96-well plates U bottom (cell culture)	Greiner	650180
Automated Metaphase Finder System Metafer4	MetaSystems	N/A
Biophotometer	Eppendorf	6133000001
Cell sorter	BD Biosciences	Aria
Centrifuge	Eppendorf	5810R
Cover glass	VWR	631-0149
Cuvettes	Roth	XK20.1
Developer machine	Protec	OPTIMAX 2010
Films	VWR	28906839
Flow cytometer	BD Biosciences	LSR Fortessa

Flow Cytometry tubes	Starstedt	55.1579
Flow Cytometry tubes (filter cap)	Falcon	352235
HiSeq 4000	Illumina	N/A
LSM700 laser scanning confocal microscope	Zeiss	N/A
MACS MS columns	Miltenyi Biotech	130-042-201
Microscope slides	Corning	2947
Mini trans blot cell	Bio-Rad	153BR107517
Nanodrop	Peqlab	ND1000
Neon Transfection System	ThermoFisher Scientific	MPK5000
Neubauer chamber	Marienfeld	0640010
NextSeq High Output 75 SR	Illumina	N/A
Novex Xcell sure lock	ThermoFisher Scientific	EI0001
OctoMACS manual separator	Miltenyi Biotech	130-042-109
Pasteur pipette	Sigma-Aldrich	7331759
PCR Machine	Applied Biosystems	4484073
Phase Lock Gel tubes	VWR	2302820
Planetary Ball Mill PM 100	Retsch	N/A
Qubit	ThermoFisher Scientific	Q33238
S220 Focused Ultrasonicator	Covaris	N/A
SpectraMAX 340PC microplate reader	Marshall Scientific	N/A
StepOnePlus Real-Time PCR System	Applied Biosystems	4376600
Thermomixer	Eppendorf	5382000015
Plasmids		
Construct	Company/Reference	Cat. No.
pMX-IRES-GFP	200	N/A
pMX-AIDER-IRES-GFP	Gift from M. Nussenzweig lab	N/A
pX458 (pSpCas9(BB)-2A-GFP)	Addgene	48138
pX330 expressing GFP	Gift from K. Rajewsky lab	Modified version of 42230
pX458 expressing $Cas9^{D10A}$	This study	N/A
pX330-GFP expressing $Cas9^{D10A}$	This study	N/A
Softwares and Algorithms		
Software/Algorithm	Company/Reference	Cat. No.
www.crispr.mit.edu	194	N/A
CRISPRGold	195	N/A
kallisto v0.43.0	204	N/A
R package tximport	205	N/A
DESeq2	206	N/A
MaxQuant software (version 1.2.2.5)	212	N/A
Burrows-Wheeler Alignment tool	213	N/A
HOMER ChIP-Seq program	214	N/A
R and ChIPseeker package	215	N/A

3.2 Methods

3.2.1 Mice

Rif1$^{FH/FH}$ (*Rif1*$^{FLAG\text{-}HA2/FLAG\text{-}HA2}$, *C57BL/6* background)[156] , *Cd19*$^{Cre/+}$ (*Cd19*$^{tm1(cre)Cgn,}$ C57BL/6 background)[188] and *Aicda*$^{-/-}$ (*Aicda*$^{tm1Hon/tm1Hon}$ C57BL/6 background)[53] were previously generated and described. The conditional *Zmynd8*^{F} allele presents two LoxP sites flaking exon 4 (ENSMUSE00001273891) for Cre-mediated recombination. It was generated by crossing the knockout first allele with conditional potential *Zmynd8*$^{tm1a(EUCOMM)Wtsi}$ (purchased at the European Mouse Mutant Archive, C57BL/6 background, EMMA #05720) with the *Rosa26*^{Flpo} deleter strain (purchased at The Jackson Laboratory, 129S4/Sv background [agouti]), which is a flippase-expressing mouse line. Germline transmission was confirmed by backcrossing progeny to a C57BL/6 mouse. Positive *Zmynd8*$^{F/+}$ heterozygous pups were intercrossed to get a homozygous *Zmynd8*$^{F/F}$ allele. Finally, these mice were bred with *Cd19*$^{Cre/+}$ mice to generate the *Zmynd8*$^{F/F}$ *Cd19*$^{Cre/+}$ mouse model (mixed *C57BL/6* 129S4/Sv background). The conditional *Pdap1*^{F} allele bears LoxP sites flanking Exon 2 (ENSMUST00000031627.8) and was generated by CRISPR/Cas9-mediated knock-in microinjection of single cell embryos transferred to a C57BL/6 female. Germline transmission was confirmed by backcrossing progeny to a C57BL/6 mouse. Positive *Pdap1*$^{F/+}$ heterozygous pups were intercrossed to get a homozygous *Pdap1*$^{F/F}$ allele. Finally, these mice were bred with *Cd19*$^{Cre/+}$ mice to generate the *Pdap1*$^{F/F}$*Cd19*$^{Cre/+}$ mouse model (C57BL/6 background). Mice were maintained in a specific pathogen-free (SPF) barrier facility under standardized conditions (20+/-2 °C temperature; 55%±15% humidity) on a 12 h light/12 h dark cycle. Male and female mice 7 to 31 weeks old were used in age-matched groups for the experiments. All experiments were performed in compliance with EU Directive 2010/63/EU, and in agreement with protocols approved by Landesamt für Gesundheit und Soziales (LAGeSo, Berlin).

3.2.2 Cell lines

WT CH12 (CH12F3, murine)[189] and clonal derivative cell lines (*Aicda*$^{-/-}$, purchased from Kerafast; *Rif1*$^{-/-}$, generated by Matteo Andreani; *Trp53bp1*$^{-/-}$, generated by Devakumar Sundaravinayagam; *Lig4*$^{-/-}$, previously described[190]; *Zmynd8*$^{-/-}$ and *Random* generated by Daniel B. Rosen and Verónica Delgado-Benito targeting exon 4 or exon 5 (ENSMUST00000109269.7); and *Pdap1*$^{-/-}$, generated by Verónica Delgado-Benito targeting exon 4 (ENSMUST0000031627.8)) were cultured in RPMI 1640 medium (ThermoFisher Scientific) supplemented with 10% FBS (fetal bovine serum; Sigma-

Aldrich), 10 mM HEPES, 1 mM Sodium Pyruvate, 1X Antibiotic Antimycotic, 2 mM L-Glutamine and 50 µM 2-Mercaptoethanol (ThermoFisher Scientific).

WT BOSC23 (human)[191], WT iMEFs (immortalized mouse embryonic fibroblasts)[98] and clonal derivative cell lines (*Brca1*$^{-/-}$, previously described[192] and *Zmynd8*$^{-/-}$ and *Random*, generated by Verónica Delgado-Benito) were cultured in DMEM medium supplemented with 10% FBS, 2 mM L-Glutamine, and 1X Penicillin-Streptomycin (ThermoFisher Scientific). All cells were grown in sterile incubators at 37 ºC and with 5% CO_2 levels.

3.2.3 Genotyping

Mice ear punches were incubated overnight with DirectPCR Ear Lysis reagent (VWR) containing 0.1 mg/ml of Proteinase K (Peqlab) at 55 ºC. Proteinase K was inactivated by incubating samples at 85 ºC for 45 min and 0.2 µl of sample was used for PCR. Amplification was performed with HotStarTaq DNA Polymerase (Qiagen). Primers used for *Zmynd8*^{F}, *Pdap1*^{F} and *Cd19*$^{Cre/+}$ genotyping listed in Table 1[193].

Regarding cell lines in culture, genomic DNA (gDNA) was extracted from at least 1x10^{6} pelleted cells that were subsequently incubated overnight with 0.1 mg/ml of Proteinase K diluted in proteinase K buffer (100 mM Tris pH 8, 0.2% SDS, 200 mM NaCl, 5 mM EDTA) at 55 ºC. Samples were transferred to 2 mL Phase Lock Gel tubes (VWR) and the same volume of Phenol:Chloroform:Isoamyl alcohol (Roth) was added. After centrifugation (13000 rpm for 20 min at RT), the upper aqueous phase was transferred to a new 1.5 ml tube and the same procedure was repeated once. The aqueous phase was then mixed with 1/10 volume of 3 M Sodium Acetate pH 5.2, 2.5X volumes of 100% ethanol and 1 mg/ml of glycogen (Roth). After 30 min incubation at -20 ºC, samples were centrifuged (13000 rpm for 15 min at 4 ºC) and precipitated gDNA was washed twice in 70% ethanol. Air-dried pellets were finally resuspended in of pre-warmed TE buffer (ThermoFisher Scientific). gDNA of *Zmynd8*- and *Pdap1*-KO clonal derivatives was amplified with HotStarTaq DNA Polymerase (Qiagen) by PCR. Primers are listed in table Table 1.

3.2.4 CRISPR/Cas9 loss-of-CSR screen

For the functional loss-of-CSR of potential RIF1-interacting factors, 3 - 6 gRNAs per candidate were cloned into the U6 cassette of pX458 plasmid (pSpCas9(BB)-2A-GFP, Addgene #48138). gRNAs were designed by using a bioinformatic tool available on-line developed by the Zhang's lab (www.crispr.mit.edu)[194]. To somatically target the selected putative phospho-53BP1 interactors, 3 gRNAs per candidate were cloned into the U6 cassette of a modified version of the pX330 vector that encodes for GFP (kind gift from Klaus Rajewsky's lab; T2A-GFP was cloned into pX330-U6-Chimeric_BB-Cbh-hspCas9,

Addgene #42230). gRNAs were designed with CrispRGold by Robin Graf (Klaus Rajewsky's lab) before this bioinformatic tool was published[195]. As a negative control for both loss-of-CSR screens, CH12 cells were transfected with empty pX458 or pX330 vectors or including gRNAs against random sequences not present in the mouse genome (GeCKO library[196]). gRNAs sequences are listed in Table 2. In all cases, cloning was performed following a previously described protocol from Zhang's lab[197]. Stbl3 bacteria were transformed with cloned plasmids and DNA was purified (PureLink HiPure Plasmid Midiprep Kit, ThermoFisher Scientific). All constructs were verified by sequencing and diagnostic digestion. $1x10^6$ bulk WT CH12 cells were electroporated with 10 µg of the Cas9-gRNAs expressing constructs (either individually or in a pooled equimolar format) by using the Neon Transfection System (ThermoFisher Scientific) and subsequently seeded in complete RPMI lacking antibiotics. Electroporating conditions were: 1600 volts, 20 milliseconds and 1 pulse. Following 40 h of incubation, nucleofected cells were collected, resuspended in PBS/FBS, filtered and sorted in bulk based on their GFP-positive expression. Samples were then resuspended at a concentration of $3x10^4$ cells/ml in complete RPMI additionally supplemented with gentamicin (0.01 mg/ml). After 72 h of recovery, cells were activated for CSR and IgA levels were determined by flow cytometry after two days (*see CSR ex vivo and in vitro*).

3.2.5 CRISPR/Cas9-mediated generation of KO clonal derivatives

For the generation of *Zmynd8*-KO CH12 and iMEFs clonal derivatives, either single gRNAs were cloned into pX458 (WT Cas9), or paired gRNAs were introduced into tandem U6 cassettes in a modified version of pX458 expressing a mutated form of Cas9 (Nickase Cas9, $Cas9^{D10A}$) to reduce possible off-target effects. To generate *Pdap1*-deficient CH12 single clones, 3 gRNA pairs were cloned into tandem U6 cassettes in a modified version of pX330-GFP encoding for $Cas9^{D10A}$ (*Nickase-a/c*). In all cases, cloning was performed using Gibson Assembly master mix (BioLabs). gRNAs sequences are listed in Table 2. Stbl3 bacteria were transformed with cloned plasmids and DNA was purified (PureLink HiPure Plasmid Midiprep Kit, ThermoFisher Scientific). All constructs were verified by sequencing and diagnostic digestion. Bulk CH12 and iMEFs cells were electroporated with the Cas9-gRNAs expressing constructs by using the Neon Transfection System (ThermoFisher Scientific) and subsequently seeded in supplemented RPMI (CH12) or DMEM (iMEFs) lacking antibiotics. Electroporating conditions for iMEFs were: 10 µg DNA per $5x10^5$ cells; 2500 volts, 10 milliseconds and 1 pulse. Following 40 h of incubation, electroporated cells were collected, resuspended in PBS/FBS, filtered and single cell sorted in 96-well plates based on their GFP-positive expression. CH12 clones were allowed to grow for 12 days in complete medium

supplemented with gentamicin (0.01 mg/ml) and subsequently re-arrayed in 3 replica 96-well plates: one plate was used for CSR activation and measuring the percentage of IgA-positive cells by flow cytometry; another plate was kept in culture to expand selected clones based on their viability and CSR levels determined by FACS; and the last plate was frozen at -80 ºC. Following 17 days of culture post-sorting, iMEFs clones were expanded. CH12 and iMEFs clonal derivatives were validated at the level of genomic scar and protein expression. Negative controls included WT cultures (non-electroporated), WT clonal derivatives and clones resulting from targeting CH12/iMEFs bulk cultures with random sequences not present in the mouse genome (GeCKO library[196]).

3.2.6 CRISPR/Cas9-mediated CSR

In order to induce CSR in CH12 independently of cytokine addition, Sμ and Sα regions were targeted by CRISPR/Cas9[198]. Single gRNAs complementary to these two regions were cloned in a modified version of the pX458 vector containing two tandem U6 cassettes. Control construct was generated by cloning two random sequences not present in the mouse genome on the same plasmid. CH12 cells were subsequently electroporated using the same conditions described above and the Neon Transfection System (ThermoFisher Scientific). Following 12 h of incubation, samples were collected, stained with fluorochrome labeled anti-IgA-PE antibody (1:400 v/v; Southern Biotech), washed, resuspended in PBS/FBS and acquired by flow cytometry.

3.2.7 Primary B cell isolation and culture

Primary B cells were isolated from the murine spleen, which was mashed with the plunger of a 3 ml syringe, resuspended in 10 ml of cold PBS/FBS (1X PBS– 3% FBS) and filtered through a 0.70 μM cell strainer. Following centrifugation, pellets were resuspended in 1 – 3 ml of ACK (ammonium-chloride-potassium) lysis buffer (ThermoFisher Scientific) in order to lyse the erythrocytes. The reaction was stopped after 3 min by adding 10 – 12 ml of cold PBS/FBS and samples were centrifuged at 1200 rpm for 5 min at 4 ºC. Each sample was resuspended in 900 μl of cold PBS/FBS and 100 μl of anti-CD43 (Ly-48) microbeads (Miltenyi Biotec). After incubation in a turning wheel at 4 ºC for 20 – 30 min, samples were diluted in 10 ml of cold PBS/FBS and filtered through a 0.45 μM mesh into a new tube. Centrifugation was carried out at 1200 rpm for 5 min at 4 ºC and pellets resuspended in 1 ml of cold PBS/FBS were transferred to previously equilibrated magnetic MACS MS columns that were attached to an OctoMACS manual separator (Miltenyi Biotec). Columns were washed with 1 extra ml of cold PBS/FBS and samples containing mature naïve B cells were eluted in 8 ml of cold

PBS/FBS. Following centrifugation (1200 rpm for 5 min at 4 ºC), pellets were resuspended in 10 – 12 ml of cold complete RPMI 1640 (10% FBS, 10 mM HEPES, 1mM Sodium Pyruvate, 2 mM L-Glutamine, 1X Antibiotic Antimycotic, 50 µM 2-Mercaptoethanol). An aliquot of cells was stained with Trypan Blue (ThermoFisher Scientific) and live mature naïve B-lymphocytes were manually counted with a Neubauer chamber. Cells were seeded at a concentration of 1.25 – 2.5x10^5 cells/ml in 6-well, 10 cm or 25cm plates and activated with different cytokines mixes (*see CSR ex vivo and in vitro*).

3.2.8 CSR *ex vivo* and *in vitro*

Primary B cells were cultured for 96 h with a combination of cytokines to induce CSR to different isotypes: 5 – 25 µg/ml LPS and 5 ng/mL of mouse recombinant IL-4 (Sigma-Aldrich) (IgG1); 5 – 25 µg/mL LPS (IgG3); ; 5 µg/ml LPS, 10 ng/ml BAFF (PeproTech) and 2 ng/ml TGFβ (R&D Systems) (IgG2b); 5 µg/ml LPS, 10 ng/ml BAFF, 2 ng/ml TGFβ and 1.5 ng/ml recombinant murine IL-5 (PeproTech) for (IgA). Cell suspensions were stained with fluorochrome-conjugated antibodies in PBS/FBS for 20 min at 4 ºC: anti-IgG1-APC (BD Biosciences), anti-IgG2b-PE (BioLegend), anti-IgG3-biotin and Streptavidin-APC (BD Biosciences), anti-IgA-PE (Southern Biotech) (1400 v/v). Cells were washed, resuspended in PBS/FBS and the percentage of switching cells was measured by flow cytometry.

CH12 cells were seeded at a concentration of 3 – 5x10^4 cells/ml in complete RPMI 1640 containing 1 – 5 µg/ml of anti-CD40 (BioLegend), 5 ng/ml TGFβ and 5 ng/ml IL-4 to induce CSR to IgA. Following 48 h of activation, cell suspensions were collected and stained with anti-IgA-PE (1:400 v/v) in PBS/FBS for 20 min at 4 °C. Cells were then washed, resuspended in PBS/FBS and CSR levels were assessed by flow cytometry.

3.2.9 Cell proliferation analysis and cell counting

Isolated naïve mature $CD43^-$ B-lymphocytes were resuspended in a 1X PBS solution containing 5 µM of CTV (cell trace violet) or in incomplete RPMI supplemented with 2 µM of CFSE (carboxyfluorescein succinimidyl ester) (ThermoFisher Scientific) at a concentration of 2x10^6 cells/ml (CTV) or 10x10^6 cells/ml (CFSE). Splenocytes were stained for 10 min at 37 ºC in the dark and subsequently washed with complete RPMI. Cells were seeded at a concentration of 1.25 – 2.5x10^5 cells/ml in complete RPMI and different cytokine mixes were added to induce CSR to different isotypes (*see CSR ex vivo and in vitro*). Once the staining was completed, stained samples plus an unstained control were acquired by flow cytometry to determine that 1) the staining worked and 2) all samples were stained equally. On the following days (48, 72 and 96 h post-seeding),

cells were collected, washed and resuspended in PBS/FBS. Proliferation profiles were compared by flow cytometry.

In order to track differences among samples in cell growth, isolated primary B-lymphocytes were seeded at a concentration of 1.25 – 2.5x10^5 cells/ml in complete RPMI and activated with LPS and IL-4 (IgG1). Following 48, 72 and 96 h post-activation, cells were stained with Trypan Blue and manually counted with a Neubauer chamber.

3.2.10 Immunization and blood serum collection

In order to induce a T cell-dependent immune reaction, 9-18 weeks-old mice were injected intraperitoneally with 100 µg of NP_{16}-CGG (4-Hydroxy-3-nitrophenylacetyl hapten conjugated to chicken gamma globulin, ratio 16; ThermoFisher Scientific) previously precipitated in alum[199]. The conjugation ratio refers to the molar ratio of hapten groups (NP) to protein (CGG) present, so in this case 16 NP haptens were conjugated to each CGG molecule. For NP_{16}-CGG precipitation, a PBS-alum solution (100 mg/ml) was freshly prepared by dissolving 1 g of $KAl(SO4)_2$ (Sigma-Aldrich) in 10 ml of 1X PBS. NP_{16}-CGG was resuspended in sterile 1X PBS at a concentration of 1 mg/ml and subsequently mixed at a 1:1 ratio in the alum solution. pH was adjusted to 6.5 – 7 by adding NaOH and, as a consequence, NP_{16}-CGG and alum precipitated together. This precipitate was washed 3 times with sterile 1X PBS by centrifugation at 5000 rpm for 15 sec. Pellet was resuspended in sterile 1X PBS to get a final NP_{16}-CGG concentration of 0.5 µg/µl and a final alum concentration of 50 µg/µl. 200 µl of this alum-precipitated NP_{16}-CGG solution were injected per mouse.

Blood was collected from the tail of animals on the same day before immunization and after 7 and 20 days. In order to obtain serum from blood, samples were first incubated in a shaking block heater at 37 ºC during 5 – 8 min, allowing for fast coagulation. The clot was subsequently removed by centrifugation at 4000 rpm for 10 min at 4 ºC. Supernatant was transferred to a new tube and samples were spin down one more time to remove the residual clot. Final supernatant was aliquoted and stored at -20 ºC.

3.2.11 ELISA (enzyme-linked immunosorbent assay)

To measure the levels of NP-specific IgM and IgG1 antibodies in serum, 96-well plates (high binding) were first coated with NP_{22}-BSA (kind gift from Klaus Rajewsky's lab) diluted in 1X BBS (NaCl, H_3BO_3, NaOH) at a concentration of 2 µg/ml. Following overnight incubation at 4 ºC, plates were washed thrice with TPBS (0.05% Tween in 1X PBS). To prepare standard curves, NP-IgG1 and NP-IgM stock aliquots (1 mg/ml; kind gift from Klaus Rajewsky's lab) were added to coated plates at a concentration of 20 µg/mL and serially diluted (4X) in TPBS. 1.63 µl of serum per sample were pipetted into

a well containing 80 µl of TPBS (100X dilution) and serial dilutions (8X) were performed. Plates were kept overnight at 4 ºC, washed thrice with TPBS and incubated with biotinylated anti-IgG1 or IgM biotinylated antibodies (Southern Biotech) at a final concentration of 1 µg/ml for at least 2 h at RT or overnight at 4 ºC. Wells were washed thrice with TPBS and incubated with Streptavidin conjugated to alkaline phosphatase (Streptavidin-AP; 1:6000 v/v; Sigma-Aldrich) for 2 h at RT. An additional washing was performed with 1 M diethanolamine buffer (10 mM diethanolamine, 0.5 mM $MgCl_2$, pH 9.8). Next, samples were treated with 2 mg/mL of 4-nitrophenyl phosphate (Sigma-Aldrich), which reacts with alkaline phosphatase and produces a yellow solution that absorbs light at 405 nm. Reaction was stopped by addition of 1N NaOH and samples were measured with SpectraMAX 340PC microplate reader (405 nm). Concentrations of NP-specific serum IgG1 and IgM antibodies were quantified relative to NP-binding IgG1 and IgM standards, respectively.

The same protocol was followed to measure total IgM and IgG1 levels in serum before immunization. In this case, plates were coated with goat anti-lambda and goat anti-kappa UNLB (Southern Biotech) and standard curves were prepared with mouse IgM- or IgG1-UNLB (Southern Biotech).

3.2.12 Retroviral infection and AID overexpression

In order to produce viral particles, BOSC23 packaging cells (containing *gag* and *pol*)[191] were seeded at a concentration of $2x10^5$ cells/ml in complete DMEM in 10 cm plates. On the following day, cells were transfected with pMX-IRES-GFP (empty vector)[200] or pMX-AIDER-IRES-GFP (vector containing the coding sequence for AID fused to the estrogen receptor; kind gift from Michel Nussenzweig's lab) and pCL-Eco (vector encoding for a murine-specific ecotropic envelope to the virus, so viral particles can only infect murine cells). Transfection was carried out by mixing the described vectors with FuGENE HD Reagent (Promega) in Opti-MEM medium (ThermoFisher Scientific), inducing the generation of viral particles that were released to the supernatant by proliferating BOSC23 cells. Splenocytes were isolated, activated with LPS and IL-4 (IgG1) or LPS, BAFF and TGFβ (IgG2b) and seeded at a concentration of $2.5x10^5$ cells/ml in 6-well plates. After 16 h, the supernatant of BOSC23 (which were transfected 48 h before) was collected with a syringe and filtered through a 0.40 µM filter. Polybrene (homemade) was added at a concentration of 10 µg/ml in order to reduce the charge repulsion between viral particles and the cell surface, increasing infection efficiency[200]. B cells tend to attach to the bottom of the plate and, therefore, the supernatant could be gently removed with a P1000 micropipette and saved. 2 ml of the virus-polybrene mix were added to each well and spinoculation was carried out by centrifuging plates at 2350 rpm

for 1.5 h at 37 ºC. Splenocytes were subsequently incubated with the virus for 4 h and then the previously saved B cell medium containing cytokines was added back to the wells. Another round of infection was performed the following day. In order to induce AID translocation to the nucleus, 0.05 µM of 4-hydroxytamoxifen (4-HT; Sigma-Aldrich) was added to the B cell cultures 72 h post-seeding. Cells were collected 24 h later, stained with fluorochrome conjugated antibodies diluted in PBS/FBS (1:400 v/v) to determine the surface expression of IgG2b-PE and IgG1-APC. Samples were incubated for 20 min at 4 ºC, washed and resuspended in PBS/FBS. CSR levels were assessed on GFP^+-gated populations (i.e. infected cells).

3.2.13 Analysis of B cell development

For analysis of murine B cell development, spleens were mashed, resuspended in cold PBS/FBS and filtered as described above (*see primary B cell isolation and culture*). To obtain cell suspensions from the bone marrow, femurs from mice were flushed in cold PBS/FBS and solutions were filtered (50 µM filter). Following centrifugation at 1200 rpm for 5 min at 4ºC, supernatants were removed, and cells were incubated with 1 – 3 mL of ACK lysis buffer for erythrocytes depletion. The reaction was stopped after 3 min by adding 10 – 12 ml of cold PBS/FBS and samples were spin down, resuspended in cold PBS/FBS, stained with Trypan Blue and counted with a Neubauer chamber. Next, cells were blocked with Fc-block (TruStain FcX, anti-mouse CD16/32; BioLegend) at a concentration of 1 µg per $1x10^6$ cells for 10 min at 4 °C. This reduces non-specific binding of the subsequently used antibodies to the Fc receptors expressed by different cell types in the spleen and bone marrow. Then, cells were labeled with fluorochrome conjugated antibodies to determine the surface expression of CD43, IgM (BD Biosciences), and CD19 (BioLegend) (bone marrow) and of CD3, CD19, IgM, CD23 (BioLegend), IgD and CD21/CD35 (BD Biosciences) (spleen). Antibodies were diluted in PBS/FBS (1:400 v/v) and samples were stained for 20 min at 4 °C. Cells were then washed, resuspended in PBS/FBS and analyzed.

3.2.14 SHM analysis

Cell suspensions from Peyer's patches of 30-31 (ZMYND8 project) and 24-27 (PDAP1 project) weeks-old mice were obtained following the same approach described above. After Fc-blocking, cells were labeled with anti-CD19, anti-B220/CD45R (BioLegend), anti-CD38 (ThermoFisher Scientific) and anti-Fas/CD95 (BD Biosciences) conjugated antibodies (1:400 v/v) for 20 min at 4 ºC in cold PBS/FBS. Following washing and filtering (50 µM filter), suspensions of non-GC (NGC; $CD38^+$ $Fas/CD95^-$) and GC B cells ($CD38^-$ $Fas/CD95^+$) were sorted directly into proteinase K buffer ($1-2x10^6$ NGC and $3-5x10^5$ GC

B cells per sample). Proteinase K (Peqlab) was added at a concentration of 0.2 mg/ml and samples were incubated overnight at 55 ºC. Genomic DNA was extracted with Phenol:Chloroform:Isoamyl alcohol (Roth) (*see Genotyping*). The 5' portions of J_H4 (*Igh*) and J_K5 (*Igk*) introns were amplified by PCR using Phusion High-Fidelity DNA Polymerase (ThermoFisher Scientific) using a forward primer that binds to a degenerate V_H or V_K family region and a reverse oligo complementary to the 3' intronic segment of J_H4 or J_K5. 800 bp J_H4 and 700 bp J_K5 PCR products were run in an agarose gel, extracted with NucleoSpin Gel and PCR Clean-up Purification Kit (Macherey-Nagel) and cloned into a pCR2.1 vector using the TOPO TA Cloning Kit (ThermoFisher Scientific). TOP10 bacteria were transformed, and purified DNA (NucleoSpin DNA Purification Kit, Macherey-Nagel) was sent for sanger sequencing. Obtained sequences were aligned to unmutated 3' intronic J_H4 or J_K5 regions and mutations were quantified over 510 bp downstream J_H4 and 536 bp downstream J_K5 gene segments. Primers used for SHM analysis are listed in Table 3[24,201].

3.2.15 Assessment of GC B cells and CSR *in vivo*

In order to measure GC and CSR levels *in vivo*, Peyer's patches from the murine intestine were isolated and grinded with two glass microscope slides with a back and forth motion in cold PBS/FBS. Cell suspensions were filtered (50 µM filter) and centrifuged at 1500 rpm for 10 min at 4ºC. Pellets were resuspended in 5 mL of cold PBS/FBS and cells were blocked with 5 µg of Fc-block (TruStain FcX, anti-mouse CD16/32; BioLegend) to reduce non-specific antibody binding. Following 10 min of incubation, cells were stained with fluorochrome-conjugated anti-CD19, anti-B220/CD45R (BioLegend) anti-CD38 (ThermoFisher Scientific), anti-Fas/CD95 (BD Biosciences) and anti-IgA (Southern Bioetch) (1:400 v/v) for 20 min at 4 ºC in cold PBS/FBS. Cells were washed, filtered, resuspended in cold PBS/FBS and analyzed by flow cytometry. GC B cells are $CD19^+$, $B220/CD45R^+$, $CD38^-$ and $Fas/CD95^+$. IgA levels were measured on gated GC B cells.

3.2.16 Metaphase analysis

iMEFs were seeded at a concentration of $3x10^4$ cells/ml (*WT*, *Rc* and *Zmynd8-KOs*) or $6x10^4$ cells/ml (*Brca1*$^{-/-}$) in complete DMEM. PARPi (Olaparib; Selleckchem) was added to the cultures on the following day at a final concentration of 2 µM. After 21 h of incubation, iMEFs were treated with 5 µg of Colcemid (Sigma-Aldrich) for 1 h to block cells in the metaphase stage of mitosis. Cells were collected and gently resuspended in pre-warmed 0.075M KCl at 37 ºC while being vortexed in order to avoid the formation of clumps. Cell suspensions were incubated at 37 ºC for 20 min to perform a hypotonic

shock. Samples were subsequently washed and fixed with a freshly prepared methanol/glacial-acetic acid (fixative) solution. Since cells were swollen due to the hypotonic shock and in order to prevent cell clumping, this solution was gently added while constantly shaking the tubes manually. Samples were incubated for 30 min at room temperature (RT) and washed twice, resuspended in fixative and kept at 4 ºC. Metaphase spreads were prepared by dropping fixed cells on tilted microscope slides with a pasteur pipette from at least 30 cm height distance. Slides were placed on a humidifier for 1 – 2 min, air-dried and incubated at 42 ºC for 1 h. Giemsa staining was performed by treating slides first with KaryoMAX Giemsa stain solution (ThermoFisher Scientific) for 2 min then with Gurr buffer solution (ThermoFisher Scientific) and lastly washed with ddH_2O. Drying was performed overnight at RT. Metaphases were acquired with the Automated Metaphase Finder System Metafer4 at 63X magnification. Chromosome radials were manually counted from at least 42 metaphases per sample.

3.2.17 RT-qPCR (quantitative reverse transcription PCR)

In order to determine *Aicda* mRNA, germline transcripts and 3'RR hs eRNAs levels in primary B cells, splenocytes were isolated and cultured with different cytokine mixes to induce CSR to different isotypes (*see CSR ex vivo and in vitro*). Following 48 h of incubation, samples were collected by centrifugation (1500 rpm for 5 min and 4 º C) and washed with 1X PBS. Total RNA was extracted using TRIzol (ThermoFisher Scientific) or RNeasy Mini Kit (Qiagen) according to manufacturer's instructions. RNA quality and concentration were determined using a NanoDrop spectrophotometer. When using TRIzol, genomic DNA (gDNA) was depleted by RapidOUT DNA Removal kit (ThermoFisher Scientific). 1 µg of RNA was used to synthesize cDNA with SuperScript VILO kit (ThermoFisher Scientific). Luna Universal qPCR Mastermix (SYBBR Green-based; NEB) was mixed with 6.25 – 12.5 ng (*Aicda* mRNA and GLTs) or 25 – 50 ng (3'RR hs eRNAs) of cDNA and 10 µM of specific forward and reverse primers (listed in Table 4)[53,202,203]. RT-qPCR was carried out in triplicates using a StepOnePlus Real-Time PCR System (Applied Biosystems) and non-retrotranscribed and non-template samples were used as negative controls. Amplification conditions were: 95 ºC for 2 min and 40 cycles of 95 ºC for 15 sec and 60 ºC for 1 min, followed by a melting curve). Analysis was performed with StepOne software and expression fold change was calculated based on double delta Ct value. Data were normalized to *Ubc* or *Gapdh* (*Aicda* mRNA and GLTs) or Gapdh (Aicda mRNA decay).

3.2.18 *Aicda* mRNA decay analysis

To assess *Aicda* mRNA decay, Actinomycin D (Sigma-Aldrich) was added at a final concentration of 10 µg/ml to the medium of IgG1- and IgG2b-activated splenocytes previously cultured for 48 h. Cells were treated for 0.5, 1, 1.5, 2, 3 and 4 h and subsequently collected by centrifugation (1500 rpm for 5 min and 4 º C). RNA was isolated with TRIzol according to manufacturer's instructions, gDNA was depleted by RapidOUT DNA Removal kit and cDNA was synthesized with SuperScript VILO kit (ThermoFisher Scientific). RT-qPCRwas performed as described above. Primers used are listed in Table 4[150].

3.2.19 RNA-seq and splicing analysis

The RNA-seq analysis for the ZMYND8 project was performed by Daniel B. Rosen at the Rockefeller University (RU) in New York. CH12 cells were cultured in activated (anti-CD40, TGFβ and IL-4) and un-activated conditions for 48 h. Samples were subsequently collected by centrifugation (1500 rpm for 5 min at 4 ºC) and RNA was isolated with AllPrep DNA/RNA Mini Kit (Qiagen). Ribosomal RNA was depleted using Ribo-Zero Gold rRNA Removal kit and libraries were prepared with TruSeq Stranded Total RNA Library Prep Kit (Illumina) following manufacturer's instructions. Three biological replicates per sample were run in two lanes on the same flow cell on NextSeq High Output 75 SR (Illumina). For data analysis (performed by Thiago Y. Oliveira and Joy A. Pai at RU), sequences were pseudo-aligned to an index created from the Ensembl mouse GRCm38.p5 assembly and custom annotations of *Igh* locus features. Transcript-level abundances were quantified using kallisto v0.43.0[204], and subsequently summarized to the gene-level using the R package tximport[205]. Differential gene expression analysis was performed using DESeq2[206].

Regarding the PDAP1 project, RNA-seq was performed in primary B cell cultures of 3 control and 3 *Pdap1*-deficient mice cultured in LPS and IL-4 (IgG1) and LPS, BAFF and TGFβ (IgG2b) for 48 h. Samples were subsequently collected by centrifugation (1500 rpm for 5 min at 4 ºC) and RNA was extracted with TRIzol. RNA quality and concentrations were determined using Qubit RNA HS Assay Kit (ThermoFisher Scientific). Ribosomal RNA was depleted using Ribo-Zero Gold rRNA Removal kit and libraries were prepared with TruSeq Stranded Total RNA Library Prep Kit Gold following manufacturer's instructions (Illumina). Samples' concentration and quality were determined using Qubit dsDNA HS Assay Kit (ThermoFisher Scientific) and library size was assessed with 4200 TapeStation System (Agilent). Samples were run in one lane on a flow cell of HiSeq 4000 (Illumina). Resulted data was analyzed by Robert Altwasser using the pigx-rnaseq pipeline[207]. STAR[208] mapped reads to Ensembl mouse

GRCm38.p5 assembly and HTSeq[209] was used to count transcript abundance. Differential gene expression analysis was performed using DESeq2[206].

For splicing analysis of *Aicda* and *Pdap1* genes we used the edgeR[210] package for R to determine differentially expressed exons. The annotation was provided by a filtered version of the GRCm38 gene annotation, which contained unique, Havana annotated exons.

3.2.20 Mutational analysis (MutPE-seq)

To determine the frequency of AID-induced mutations at 5' Sμ, primary B cell cultures were activated with LPS and IL-4 (IgG1) for 72 or 96 h. Samples were collected by centrifugation (1500 rpm for 5 min at 4 ºC) and gDNA was extracted with phenol-chloroform-isoamyl alcohol (Roth) (*see Genotyping*). To amplify 5' Sμ amplicons, 100 ng of gDNA were mixed with Phusion High-Fidelity DNA Polymerase (ThermoFisher Scientific or NEB) and locus-specific primers for 5' Sμ. PCR conditions were: 98 ºC for 30 sec, 20 cycles of 98 ºC for 10 sec, 64.9 ºC for 15 sec and 72 ºC for 30 sec and a final extension step of 72 ºC for 5 min. DNA was gel-extracted (250 bp) and purified by NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) and 2/3 of the resulting product were used as template for a second PCR to introduce sequencing adapters and sample-specific indexes. Amplification conditions were: 98 ºC for 30 sec, 15 cycles of 98 ºC for 10 sec, 64.6 ºC for 15 sec and 72 ºC for 30 sec and a final extension step of 72 ºC for 5 min Gel-extracted and purified amplicons (390 bp) were pooled and sequenced by MiSeq (2 x 300 bp paired-end read sequencing). Mutations present in both paired reads were considered as AID-induced mutations. Bioinformatic analysis was performed by Thiago Y. Oliveira and Joy A. Pai at RU. Primers used are listed in Table 5[93].

3.2.21 Clonogenic survival assay

To determine cell survival following DNA damage, iMEFs were plated in 60 mm dishes Following 24 h of culture, cells were either irradiated with the indicated doses or treated with 0.1, 1 and 5 μM of PARPi (Olaparib; Selleckchem). In the latter case, media and PARPi were replenished after 7 days. iMEFs were incubated for a total of 14 days post-seeding and supernatant was aspirated from the plates. 3 biological and 3 technical replicates were performed. Colonies were subsequently fixed with 15% acetic acid-methanol solution (v/v) for 5 min, stained with 0.5% Crystal Violet (Sigma-Aldrich) for 30 min and washed twice with 1X PBS. Plates were scanned and colonies were quantified with manually with ImageJ software.

3.2.22 Switch junction analysis

To amplify Sμ – Sγ1 junctions, primary B cells were isolated and activated with LPS and IL-4 (IgG1). Following 96 h of culturing, gDNA was extracted using phenol:chloroform:isoamyl alcohol (Roth) (*see Genotyping*). Amplification was carried out by two rounds of PCR using Phusion High-Fidelity DNA Polymerase (ThermoFisher Scientific). Amplification conditions for the first PCR were: 98 ºC for 30 sec, 10 cycles of 98 ºC for 10 sec and 72 ºC for 2 min and a final extension step of 72 ºC for 10 min. Amplification conditions for the second PCR were: 98 ºC for 30 sec, 20 cycles of 98 ºC for 10 sec and 72 ºC for 2 min and a final extension step of 72 ºC for 10 min. Amplicons were run on agarose gels and 350 – 1000 bp fragments were extracted with NucleoSpin Gel and PCR Clean-up Purification Kit (Macherey-Nagel) and cloned into a pCR2.1 vector using the TOPO TA Cloning Kit (ThermoFisher Scientific). TOP10 bacteria were transformed, and purified DNA (NucleoSpin DNA Purification Kit, Macherey-Nagel) was sent for sanger sequencing with M13 forward and reverse universal oligos. Primers used for this analysis are listed in Table 6[96].

3.2.23 Cell lysates and immunoblotting

Frozen cell pellets at -80 ºC were thawed by resuspension in cold lysis buffer (20X complete protease inhibitor and 1000X 1 M DTT (ThermoFisher Scientific) in RIPA buffer (Sigma-Aldrich)). Samples were incubated for 20 min at 4 ºC and subsequently centrifuged at 12700 rpm for 10 min at 4 ºC. Supernatant was used to determine protein concentration by Bradford Assay. Standard curve was prepared with different concentrations of BSA (10 mg/ml) diluted in lysis buffer. Lysates were mixed with 4X NuPAGE LDS Protein Loading buffer (ThermoFisher Scientific) supplemented with 6X 1 M DTT and boiled at 72 ºC for 10 min. Samples were loaded into precast polyacrylamide gradient gels (NuPAGE 3-8% Tris-Acetate or 4-12% Bis-Tris; ThermoFisher Scientific) and proteins were separated by electrophoresis in running buffer (NuPAGE MES, MOPS or Tris-Acetate SDS) at 120 V until the dye reached the front. Proteins were transferred to a PVDF (polyvinylidene difluoride, Millipore) membrane using the Mini Trans-Blot Tank (ThermoFisher Scientific) containing transfer buffer (25nM Tris, 192 mM glycine, 20% methanol) at 200 mA for 1.5 – 2 h. To reduce unspecific antibody binding, membranes were incubated in blocking buffer (3% BSA (Albumin IgG-free; Roth) in 1X PBS – 0.1% Tween) shaking for 1 h at RT or overnight at 4 ºC. Primary antibodies were diluted in blocking buffer and added to membranes that were incubated in sealed pouches for 1 h at RT or overnight at 4 ºC (constant shaking): rabbit anti-ZMYND8, mouse anti-βActin, anti-PDAP1 (Sigma-Aldrich), rat anti-AID (ThermoFisher Scientific). Membranes were subsequently washed 3 times in 1X PBS – 0.1% Tween for 10 min at RT and constant

agitation. Horseradish peroxidase (HRP)-tagged secondary antibodies raised against mouse, rabbit or rat (kind gift from Fritz Rathjen's lab) IgG were diluted in blocking buffer (1:20000, v/v) and membranes were incubated for 1 h at RT (constant shaking). Following 3 additional washes in 1X PBS – 0.1% Tween, membranes were activated with freshly prepared chemiluminescence solution (PerkinElmer) and exposed to films (Amersham hyperfilm ECL, VWR) that were developed using the OPTIMAX 2010 film developer.

3.2.24 I-DIRT and immunoisolation of RIF1^{FH} complexes

Primary B cells were isolated from the spleen of *Rif1*$^{FH/FH}$ and *wt* mice as described above. Splenocytes were cultured in complete SILAC-RPMI (-Arg, -Lys) medium (ThermoFisher Scientific) supplemented dialized FBS (ThermoFisher Scientific) and with either $^{13}C_6$ L- arginine and $^{13}C_6$ L- lysine (heavy medium; *Rif1*$^{FH/FH}$; Cambridge Isotope Laboratories) or non-labeled L-arginine and L-lysine (light medium; *wt*; Sigma-Aldrich), respectively. Cells were treated with LPS and IL-4 to induce CSR to IgG1 and also with anti-CD180 (RP/14; BD Biosciences) to promote B cell proliferation. Following 96 h of incubation, splenocytes were irradiated (20 Gy) to increase the amount of DSBs and put back in culture for 45 min to allow for recovery. 2 x10^9 cells per genotype were collected by centrifugation, resuspended in 20 mM HEPES containing 1.2% polyvinylpyrrolidone (Sigma-Aldrich), protease and phosphatase inhibitor cocktails (Roche), 0.5 mM DTT, and frozen in liquid nitrogen. *Rif1*$^{FH/FH}$ and *wt* frozen cells were mixed in a 1:1 ratio and cryogenically lysed by wet grinding in a Planetary Ball Mill PM 100. The resulting frozen product was thawed by resuspension in extraction buffer (20 mM Tris-Cl pH 8, 150 nM NaCl, 0.5% Igepal CA-630 (Sigma-Aldrich), 1.5 mM $MgCl_2$, Benzonase (Sigma-Aldrich) and protease and phosphatase inhibitor cocktails). Glutaraldehyde (Sigma-Aldrich) was added to the buffer (2.5 mM) in order to stabilize labile interactions while preserving the native composition of protein complexes[211]. Following 5 min incubation, reaction was stopped with 100 mM Tris-Cl pH 8.0 and lysate was centrifuged at 13000 rpm at 4 ºC. Supernatant was directly incubated with magnetic beads (M-270 Epoxy beads; ThermoFisher Scientific) conjugated to anti-Flag M2 antibody (Sigma-Aldrich) for 1 h at 4 ºC[96]. Sample was washed with extraction buffer and RIF1^{FH} bait and associated protein were eluted twice under native conditions by two rounds of incubation with 2.5 µg/µl 3XFlag peptide (Sigma-Aldrich). Following 45 min of incubation at 4 ºC (constant shaking), sample was loaded into a 4 – 12% Bis-Tris precast polyacrylamide gradient gel (ThermoFisher Scientific). Proteins were separated by electrophoresis and subsequently stained with Coomassie blue. The gel was divided into upper and lower parts based on the 39 kDa molecular weight marker. The upper part

was cut in 5 portions and all gel pieces were subjected to in-gel tryptic digestion. Peptides were extracted and purified, analyzed by LCMS using a Thermo Q Exactive Plus mass spectrometer, with a Thermo Easy-nLC 1000 HPLC and a Thermo Easy-Spray electrospray source. Identification and quantification of proteins was performed by searching against a mouse protein sequence database with the MaxQuant software (version 1.2.2.5)[212]. Dan B. Rosen performed the I-DIRT, immunoisolation of RIF1-containing complexes and mass spectrometry in collaboration with the laboratory of Brian T. Chait at the RU.

3.2.25 Immunofluorescence

iMEFs were seeded on coverslips placed into 12-well plates that were incubated overnight. Cultures were subsequently irradiated (10 Gy) and put back in culture for either 90 min or 6 h to allow for recovery. Cells were gently washed with 1X PBS and fixed with 2 ml of freshly prepared 4% paraformaldehyde solution in 1X PBS (Sigma-Aldrich) for 10 min at RT. After three washings, permeabilization was performed with 2 ml of 0.5% Triton X-100 for 5 min at RT. Plates were additionally rinsed 3 times with 1X PBS and incubated with blocking solution (2.5% goat serum and 1% BSA in 1X PBS) shaking for 1 h at RT or overnight at 4 ºC. Primary antibodies were diluted in blocking solution and cells were stained with mouse anti-γH2AX (1:500 v/v; Millipore), rabbit anti-RIF1 serum[96], mouse anti-53BP1 (Millipore), mouse anti-Flag M2 or rabbit anti-ZMYND8 (Sigma-Aldrich) for 1 h at RT. Coverslips were washed once with 2 ml of 0.1% Triton X-100 (in 1X PBS) for 5 min at RT and subsequently rinsed twice with 1X PBS. Samples were incubated with fluorochrome-conjugated secondary antibodies goat anti-rabbit Alexa546 and goat anti-mouse Alexa488 (1:500 v/v; ThermoFisher Scientific) together with DAPI (1:1000 v/v) for 1 h at RT. Coverslips were rinsed three times with 1X PBS and placed on microscope slides using mounting medium (Prolong Gold Antifade Mountant, ThermoFisher Scientific). Images were acquired with an inverted LSM700 laser scanning confocal microscope (Zeiss) and processed with Fiji software. These experiments and analyses were performed by Devakumar Sundaravinayagam.

3.2.26 ChIP-seq (chromatin immunoprecipitation coupled with sequencing)

ChIP-seq analysis for ZMYND8 in CH12 was performed by Qiao Wang. ChIP-seq analysis for ZMYND8 and RNAPII in primary B cells were conducted by me (splenocytes isolation, culture and fixing) and Daniel B. Rosen (libraries preparation). Computational analysis was done by Thiago Y. Oliveira and Joy A. Pai. Primary B cells were extracted and stimulated with LPS and IL-4 for 72 h. CH12 cells were cultured in activating (anti-CD40, TGFβ and IL-4) or non-activating conditions for 48 h. Samples were counted,

collected by centrifugation (1500 rpm for 5 min at 4 ºC) and fixed with 1% of 16% Formaldehyde Methanol-free (ThermoFisher Scientific) for 10 min at 37 ºC. Reaction was quenched by adding 1/20 volume of 2.5M glycine dissolved in 1X PBS pH 7.4 (final concentration 0.125 M) and swirling. Fixed cells were washed with cold 1X PBS, centrifuged (1500 rpm for 5 min at 4 ºC), aliquoted (20 x10^6 cells/aliquot), frozen in liquid nitrogen and stored at -80 ºC. Frozen pellets were resuspended in 100 µl of 1% SDS of RIPA buffer supplemented with Complete-EDTA free protease inhibitor (Roche) and incubated for 20 min at 4 ºC. Sonication was performed with a Covaris S220 Focused Ultrasonicator at peak value 105, duty factor 5, cycle/burst 200 for 10 min. Next, SDS concentration was adjusted to 0.1% by diluting samples with non-SDS containing RIPA buffer and tubes were centrifuged at 14000 rpm for 10 min at 4 ºC. Chromatin fragments were pre-cleared by incubation with Dynabeads Protein A (ThermoFisher Scientific) with rotation at 4 ºC for 1 h. Samples were subsequently mixed with 10 µg of antibody specific for ZMYND8 (Sigma-Aldrich) or RNAPII (phospho-S5; 4H8, Abcam) previously bound to Dynabeads. This allowed for immunoprecipitation of cross-linked DNA fragments associated with either ZMYND8 or RNAPII. Samples were incubated with Proteinase K (Peqlab) for 4 h at 65 ºC and DNA was extracted with phenol:chloroform:isoamyl alcohol (Roth). DNA libraries were prepared using Illumina compatible adaptors (Bio Scientific) and sequenced on an Illumina NextSeq 500 sequencer (Illumina). FASTQ files were aligned against mouse genome (mm10) using Burrows-Wheeler Alignment tool[213]. Processing and peak-calling of ChIP-seq data were accomplished with HOMER ChIP-seq program[214]. Peak annotation was done using the ChIPseeker R package[215].

4. RESULTS

4.1 Identification of novel CSR factors

4.1.1 Functional loss-of-CSR screen

To identify *bona-fide* effectors involved in CSR, a robust functional screen was established in the B cell lymphoma line CH12F13 (CH12) that can be induced to express AID and undergo class switching to IgA with high efficiency upon *in vitro* cytokine treatment[189,216]. B cells cannot be transfected *via* classical methods (calcium phosphate, lipofectamin, etc.), and the efficiency of retroviral transduction is drastically affected by the size of the Cas9-containing vectors (>10 Kb). Therefore, as a first step, the conditions for efficient expression of Cas9 and guide RNAs (gRNAs) targeting were set up by electroporating a pX330 variant vector that contains GFP co-translationally expressed with Cas9 (pX330-T2A-GFP)[197,217]. Different electroporation parameters (voltage, width and pulse) were tested in order to find the optimal conditions for efficient transfection of CH12 cells, being 1600 volts, 20 milliseconds and 1 pulse the best combination (data not shown). Subsequently, the loss-of-CSR assay was set up by targeting *Aicda* gene, which encodes for AID and it is an essential factor for antibody isotype differentiation[53]. Positively nucleofected cells somatically targeted with 3 independent gRNAs against *Aicda* showed an 80-90% decrease in CSR when compared to control (cells electroporated with an empty pX330-T2A-GFP vector, EV) (**Figure 4.1A**). Next, several gRNAs were designed to disrupt *Trp53bp1* gene that encode for the DSB repair factor 53BP1, since 53BP1-deficient primary mouse B cells exhibit a dramatic reduction of CSR efficiency (more than 90%)[218,219]. Surprisingly, no major reduction in the percentage of IgA positive cells was observed when employing the same conditions that resulted in a drastic CSR deficiency of *Aicda*-targeted CH12 cells (**Figure 4.1B, left panel**). This could be due to a protein turnover issue, since AID expression is induced only upon activation for CSR and 53BP1 is constitutively expressed. Therefore, even if *Trp53bp1* had been correctly disrupted following Cas9-gRNA electroporation, it is possible that the previously translated protein was still available when cells were stimulated to undergo CSR. To test this, the time in culture in between nucleofection and cytokine treatment was extended by sorting the positively electroporated GFP positive cells, allowing the cells to undergo more cell divisions and 53BP1 proteasomal degradation. Accordingly, the residual CSR levels dropped to 40-50%, which ensured a large dynamic range to further perform the screen (**Figure 4.1B, right panel**). These experimental conditions (**Figure 4.1C**) were further validated by targeting *Rif1*, since deletion of this factor in primary B cells leads to a 80-85% decrease in CSR[96]. In line with *Trp53bp1* disruption in CH12 and the fact that

53BP1-deficiency affects CSR more dramatically than Rif1 absence in primary B cell cultures, *Rif1*-targeted cells showed a 50-60% reduction in the percentage of IgA positive cells (**Figure 1D**). Altogether, a robust loss-of-CSR assay was set up for the functional screening of potential candidates involved in CSR.

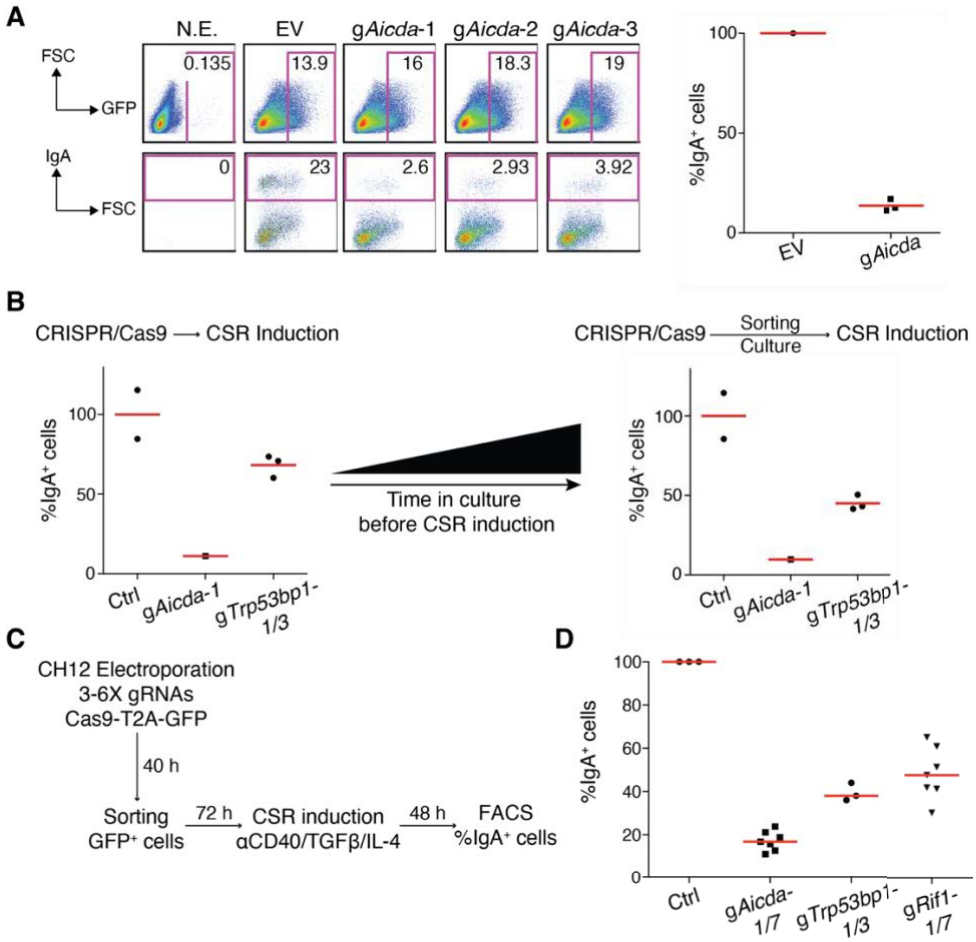

Figure 4.1. Loss-of-CSR CRISPR/Cas9 in CH12 cells.
A. Left: flow cytometry plots measuring the percentage of nucleofected CH12 cells (GFP$^+$; top) and CSR to IgA (bottom) after CRISPR/Cas9 somatic targeting of *Aicda* with three different gRNAs. Right: summary graph representing CSR to IgA relative to the empty vector (EV) value. N.E., non-electroporated. **B.** Dot plots showing CSR to IgA in Cas9/g*Aicda*- and Cas9/gTrp*53bp1*-nucleofected CH12 cells activated subsequently to electroporation (left) or after sorting and culturing the positively targeted cells (right). Data were normalized to the mean value of the control samples (Ctrl: EV or g*Random*), which was set to 100%. **C.** Scheme representing the experimental conditions for the loss-of-CSR screen by CRISPR/Cas9 in CH12 cells. **D.** Graph summarizing the validation of the loss-of-CSR screen in CH12 by somatic targeting of *Aicda*, Trp*53bp1* and *Rif1* genes following the scheme in (C). Graph summarizes three independent experiments where either Aicda, Rif1 or Trp53bp1 was targeted with different gRNAs. CSR is represented as a percentage of the control (EV or g*Random*) within the same experiment.

4.1.2 Identification of potential factors involved in CSR

The RIF1 protein interactome in primary B-lymphocytes undergoing CSR was studied in order to define the pool of novel factors that could be regulating antibody isotype differentiation. A SILAC-based (Stable Isotype Labeling by Amino acids in Cell culture) approach was optimized for the targeted identification of *bona-fide,* and likely transient, RIF1 interactors (Isotopic Differentiation of Interactors as Random or Targeted, I-DIRT)[220]. This method consists on culturing cells expressing a tagged protein of interest and un-tagged cells in media that are differentially isotopically labeled, allowing for identification of specific interactors of the tagged factor. For this purpose, mice bearing a homozygous 1X Flag – 2X hemagglutinin (HA) – tagged knocked-in version of RIF1 that can still support CSR to wild-type (WT) levels were used (*Rif1*$^{FH/FH}$). WT cells expressing endogenous RIF1 were isolated in parallel and cultured in isotopically labeled light medium whereas RIF1$_{FLAG\text{-}HA}$ splenocytes were cultured in isotopically labeled heavy medium. Both cell populations were irradiated (20 Gy) to increase DSB formation and processed for anti-FLAG immunoprecipitation (**Figure 4.2A)**. A collaboration was established with Brian Chait's laboratory at the Rockefeller University (RU) to optimize the experimental conditions in order to identify RIF1 weak, transient or rapidly exchanging interactions[221]. Specifically, cells were frozen in liquid nitrogen to preserve native interactions and subsequently subjected to cryomilling to generate frozen cell micron fragments that can be rapidly deepen in a solvent. Lastly, glutaraldehyde was added to chemically stabilize protein complexes before cell lysis. This cross-linking reagent had been previously shown to have a high reactivity and stabilizing properties at low temperatures[222,223].

Upon immuno-isolation of RIF1-interacting complexes, the vast majority of identified proteins displayed a SILAC ratio characteristic of pull-down contaminant, with several proteins exhibiting a ratio significantly higher than the mean of the distribution, being RIF1 bait and 53BP1 the top hits (**Figure 4.2B**). To increase the confidence selecting potential new RIF1-interacting candidates, the subsequent functional loss-of-CSR analysis was limited to candidates that: 1) exhibited a SILAC ratio more than 2 standard deviations (SD) (0.095) above the mean of the distribution (0.49); 2) have been identified with at least 4 peptides; and 3) exhibited an overall PEP (Posterior Error Probability)[224] $\leq 1x10^{-4}$. Moreover, the resulting candidates were shortlisted based on previous literature suggesting a role in DNA repair, leading to a final list of fourteen potential RIF1 interactors. The loss-of-CSR assay was performed on these factors and IgA levels were measured after targeting each of them with three to six different pooled gRNAs. "Random" gRNAs (sequences not present in the mouse genome), and gRNAs against *Aicda*, *Rif1* and *Trp53bp* were used in parallel as negative and positive controls for the

assay, respectively. gRNAs were designed using a robust online tool (www.crispr.mit.edu)[194] that predicted sequence-specific gRNAs and off-target effects. For each candidate, a literature search was done in order to look for available KO (knock-out) mouse models or cell lines that gave information about which exon to target. ZMYND8 (Zinc Finger MYND-type [myeloid, Nervy, and DEAF-1] containing 8), also known as RACK7 and PRKCBP1, was one of the top potential RIF1 interacting candidates and was identified as a hit. Consequently, its genomic disruption led to a clear CSR decrease in CH12 (**Figure 4.2C**). This factor had been previously shown to be a chromatin reader that binds to enhancers and super-enhancers to regulate transcription, acting mostly as a repressor[225–228]. Moreover, it had been described to act at sites of DNA damaged within transcriptionally active chromatin to repress transcription and promote HR-dependent repair[229–233]. Considering that transcription, DNA repair and enhancer activity are crucial steps in CSR, ZMYND8 was selected for further studying its role in regulating antibody isotype differentiation.

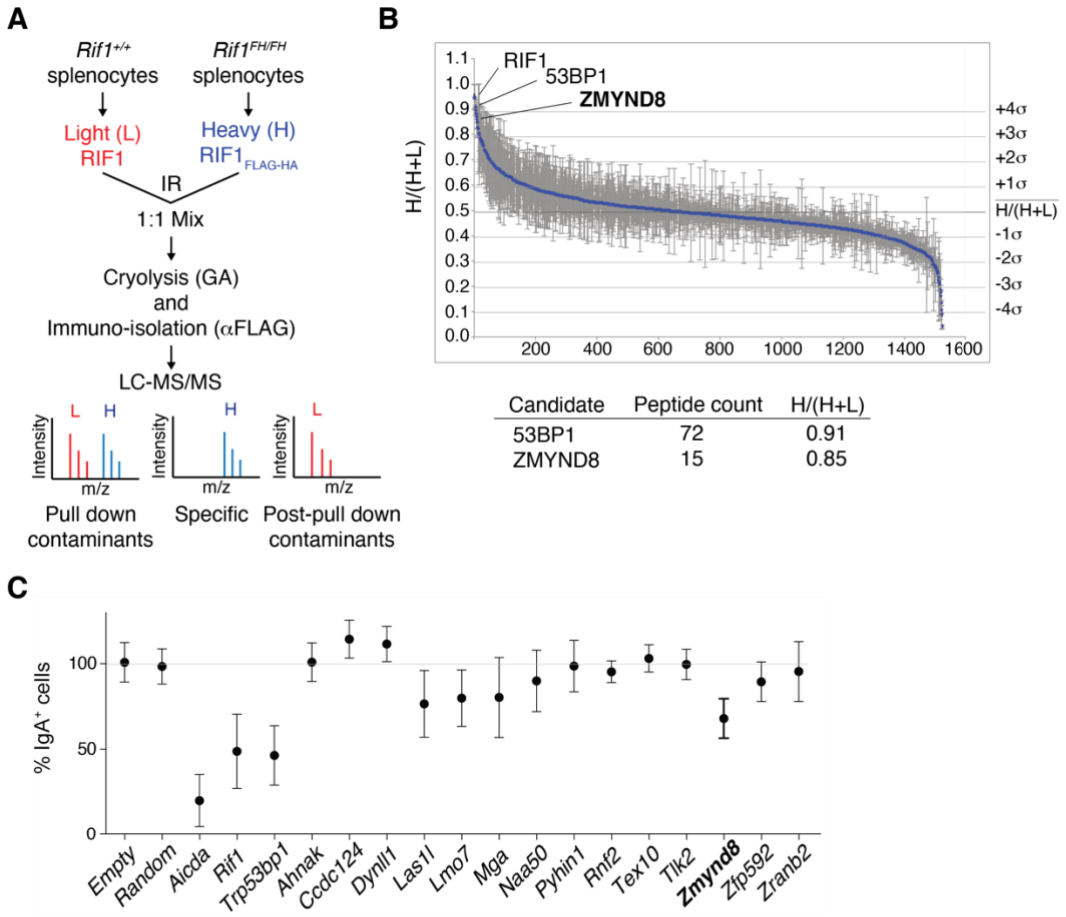

Figure 4.2. Screening for novel RIF1-interacting CSR factors in B-lymphocytes.
A. Experimental scheme of the identification of RIF1-interactome in stimulated primary B-lymphocytes by I-DIRT (Isotopic Differentiation of Interactions as Random or Targeted). FH: FLAG-HA tag; IR: ionizing radiation; GA: glutaraldehyde; LC-MS/MS: liquid chromatography-tandem mass spectrometry. **B.** Top: graph depicting the RIF1-interacting proteins identified by I-

DIRT following the scheme in (A) and plotted based on their H/(H+L) ratio. Data was filtered to show factors with a peptide count ≥ 4 and a posterior error probability ≤ 10^{-4}. Each candidate's error bar indicates the standard error of its H/(H+L) mean value. Bottom: table specifying the peptide count and H/(H+L) mean for 53BP1 and ZMYND8. **C.** Graph showing CSR to IgA in activated CH12 cells after CRISPR/Cas9 somatic targeting of the indicated candidates resulting from the RIF1 I-DIRT. At least three independent experiments were performed for each candidate. Error bars indicate SD.

An independent list of potential phospho-dependent 53BP1 interactors identified by SILAC-based approach that was previously published was used for the functional screen of loss-of-CSR[96]. Briefly, *Trp53bp1*-deficient primary B cultures undergoing CSR were infected with a retrovirus containing either the cDNA for a WT or for a phospho-deficient version of 53BP1 fused to a FLAG tag (53BP1$^{PM\text{-}FLAG}$). After isotopic labeling, cells were irradiated to increase the load of DSBs, and 53BP1-interacting factors were immunoprecipitated with an anti-FLAG antibody (**Figure 4.3A**). Since 53BP1 depends on ATM-mediated phosphorylation to function in CSR during DSB repair[98,99,147,148], potential candidates that specifically interacted with WT 53BP1 (that can undergo phosphorylation) were selected. The SD, peptide number and PEP cut-off values described above were also employed here to increase the confidence in the selection of potential interactors. In this case, a list of thirty-three potential phospho-53BP1 interactors were assessed for their function in CSR. In this case, gRNAs were designed with CrispRGold, which predicts off-target effects and provides sequences that hit the maximum number of isoforms[195]. Upon genomic targeting by CRISPR/Cas9, *Mrpl14*, *Eif4g2* and *Pdap1* showed the most dramatic reduction in IgA levels compared to "*Random*" controls (**Figure 4.3B**). MRPL14 is a component of the mitochondrial large ribosomal subunit[234–236], which seemed to be important for B cell viability (data not shown). EIF4G2 (eukaryotic translation initiation factor 4 gamma 2) is a cap-binding factor that plays a role in translation initiation[237–239]. PDAP1 (platelet-derived growth factor-associated protein 1) is a protein of 181 amino acids highly conserved in vertebrates initially identified as a casein kinase II substrate[240,241]. It has been described to putatively bind RNA in human cell lines in several RNA-protein interactome studies[242–246]. An additional study showed that PDAP1 is overexpressed in glioma cells and it is related to malignant proliferation through up-regulation of PDGF-B/Akt/PDK1 signaling[247]. Lastly, a recent publication suggested a potential role of PDAP1 in STAT6 – ERK – NF-κB network[248]. Given the limited information available in literature about the function in PDAP1 and its degree of conservation, this candidate was selected for further studying its role in regulating antibody isotype differentiation.

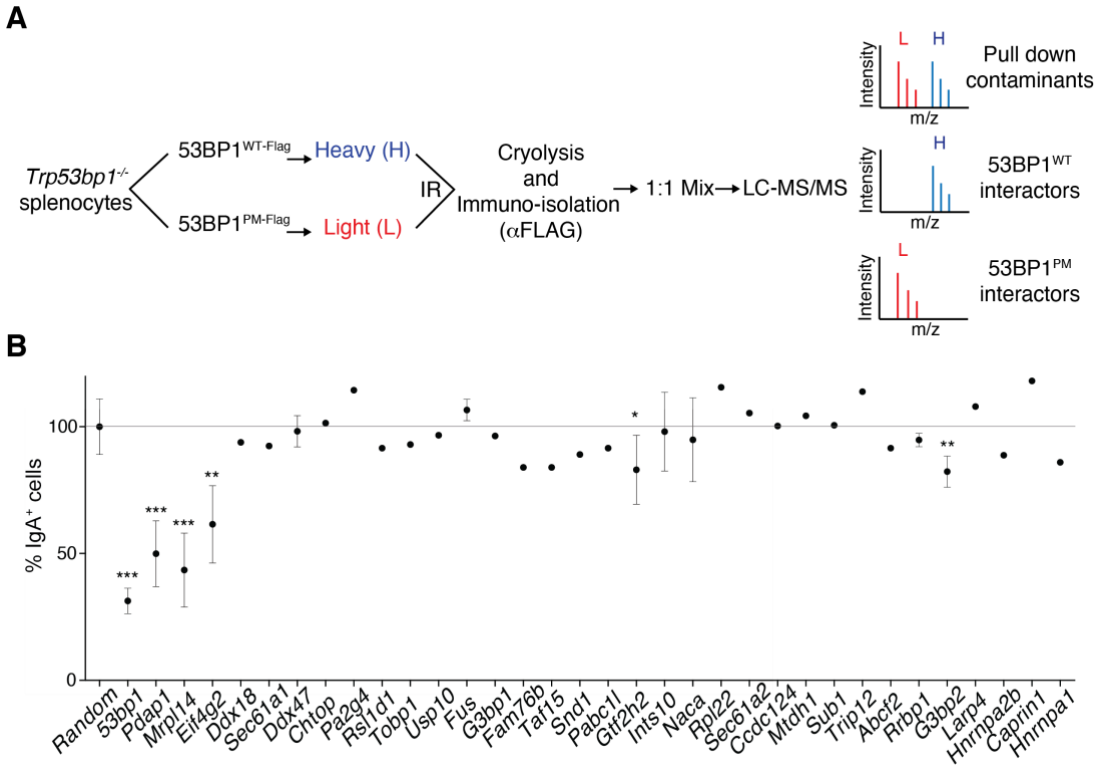

Figure 4.3. Screening for novel phospho-53BP1-interacting CSR factors in B-lymphocytes. **A.** Experimental scheme of the identification of 53BP1-interactome in stimulated primary B-lymphocytes by SILAC (Stable Isotope Labeling by Amino acids in Cell culture)[96]. 53BP1$^{WT\text{-}FLAG}$: 53BP1$^{Wild\text{-}Type\text{-}FLAG}$; 53BP1$^{PM\text{-}FLAG}$: 53BP1$^{Phospho\text{-}Mutant\text{-}FLAG}$. LC-MS/MS: liquid chromatography-tandem mass spectrometry. IR: ionized radiation. **B.** Graph showing the percentage of IgA positive cells in activated CH12 after CRISPR/Cas9 somatic targeting of the indicated candidates resulting from the 53BP1 SILAC. At least three independent experiments were performed for each candidate where error bars depicting the SD are shown. Significance was calculated with the Mann-Whitney U test. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.

4.2 The chromatin reader ZMYND8 regulates the 3' *Igh* enhancer to promote CSR

4.2.1 ZMYND8 is essential for efficient CSR

In agreement with the loss-of-CSR screen results showed above, somatic targeting of *Zmynd8* in CH12 cells led to a significant decrease in the percentage of IgA positive cells and ZMYND8 protein expression, compared to control (cells electroporated with an EV or with a "random" gRNA that does not target any region in the mouse genome) (**Figure 4.4A-B**). To further dissect the role of this factor in CSR, *Zmynd8*- KO clonal derivatives were generated by using the nickase version of Cas9 (Nickase-Cas9^{D10A}), which exhibits reduced off-target effects[197]. This mutant version of the protein generates a nick instead of a DSB. Therefore, only when using two gRNAs that target close regions in opposite strands, the two nicks are converted into a DSB. As expected, these KO clones were defective in CSR and lacked ZMYND8 protein expression (**Figure 4.4C**).

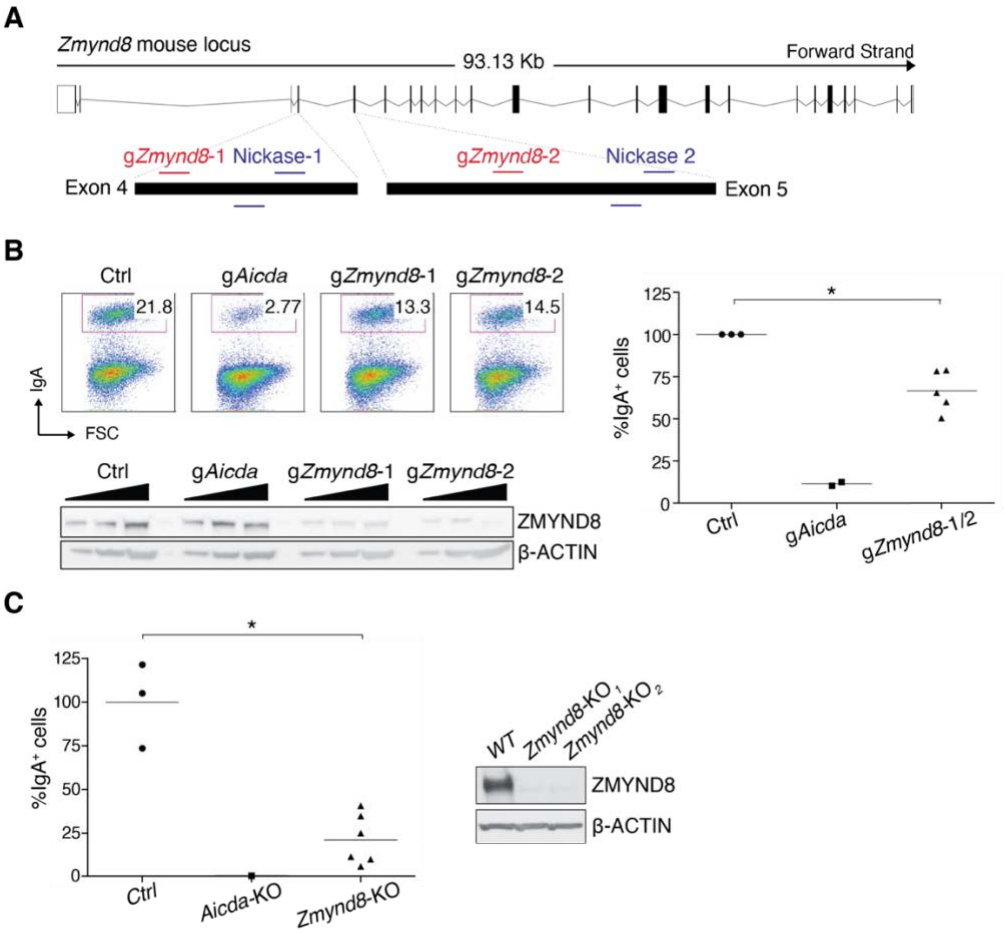

Figure 4.4. Deletion of *Zmynd8* in CH12 causes loss of CSR.
A. Representation of the murine *Zmynd8* locus and the genomic location of the gRNAs used in this study. Single gRNAs that guide a WT-Cas9 are shown in red, whereas gRNA pairs that guide a Nickase-Cas9^{D10A} are depicted in blue. **B.** Flow cytometry plots representing the percentage of IgA positive cells (left top) and Western Blot (WB) analysis of WCE (whole-cell-extracts) (left bottom) from somatically targeted CH12 cells with CRISPR/Cas9 and the indicated gRNAs. Ctrl (control) stands for cells electroporated with an EV or with a g*Random*. Right: graph summarizing the CSR data to IgA of three independent experiments after normalizing data to the control value that is set to 100%. **C.** Left: summary dot plot of IgA levels in *Random Aicda*- and *Zmynd8*-KO activated CH12 cells (clonal derivatives). Values of control cell lines were averaged and set to 100% for normalization. Data is representative of two independent experiments. Right: WB analysis of WCE from WT CH12 cells and two independent *Zmynd8*-KO clonal derivatives. Significance in (B) and (C) was calculated with the Mann-Whitney U test. *p ≤ 0.05.

To confirm the results in CH12 and dive into the mechanistic aspects of ZMYND8 in antibody isotype differentiation, a conditional *Zmynd8*-KO mouse model was generated. Specifically, the KO-first allele with conditional potential *Zmynd8*$^{tm1a(EUCOMM)Wtsi}$ mouse was purchased and crossed to a flippase-expressing mouse line (*Rosa26*^{Flpo}). Next, *Zmynd8*$^{F/+}$ mice, exhibiting two LoxP sites flanking exon 4, were backcrossed to a C57BL/6 (*wt*) mouse to confirm germline transmission. Positive *Zmynd8*$^{F/+}$ heterozygous pups were intercrossed to get a homozygous *Zmynd8*$^{F/F}$ allele. Finally, these mice were crossed to a *Cd19*$^{Cre/+}$ mouse to conditionally delete *Zmynd8* in B-lymphocytes

(*Zmynd8*$^{F/F}$ *Cd19*$^{Cre/+}$) (**Figure 4.5A**). Of note, CD19 is a B cell marker that is expressed early on during development, specifically at the pro-B cell stage (*see section 1.3*)[37,188]. This implies that *Zmynd8* is deleted at the initial stages of B cell differentiation when expressing Cre under the control of CD19. As expected, primary B-lymphocytes from the newly generated mouse model *Zmynd8*$^{F/F}$-*Cd19*$^{Cre/+}$ showed a major reduction of ZMYND8 protein expression as determined by WB analysis (**Figure 4.5B**).

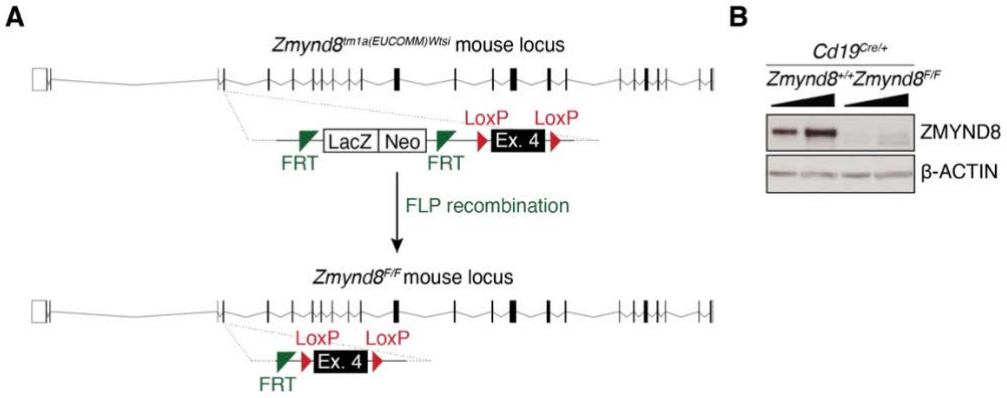

Figure 4.5. Generation of a conditional *Zmynd8* knock-out mouse model.
A. Top: representation of the KO-first allele with conditional potential *Zmynd8*^{tm1a} mouse model. Bottom: scheme of the same locus after flippase (FLP) recombination. FRT: flippase recognition target; LoxP: lox site. **B.** WB analysis of WCE from splenocytes of the indicated genotypes cultured for 72 h in the presence of LPS and IL-4. Triangles represent threefold dilution.

B cells present different markers on their surface depending on their developmental stage, which can be used as a tool to determine distinctive B cell populations by flow cytometry (*see section 1.3*)[36]. This approach was used to study B cell development in ZMYND8-deficient cells collected from the spleen and the bone marrow (BM). B-lymphocytes seemed to differentiate comparably to control, evidencing that ZMYND8 does not play a role in B cell development or V(D)J recombination (**Figure 4.6A-B**). Moreover, ZMYND8 protein expression was not induced upon cytokine stimulation and levels did not vary during CSR induction as assessed by WB analysis of *Cd19*$^{Cre/+}$ splenocytes (**Figure 4.6C**).

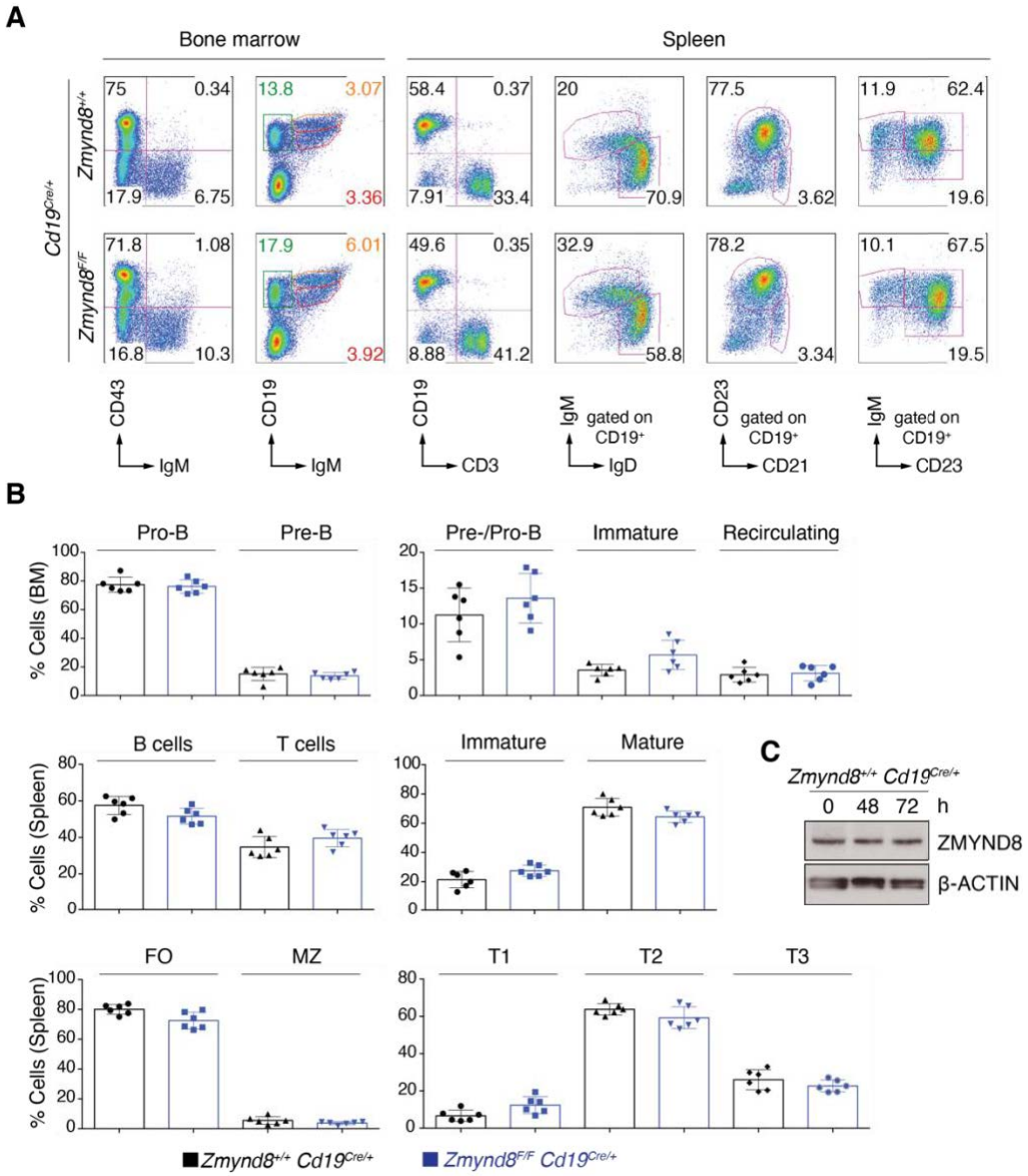

Figure 4.6. ZMYND8 is dispensable for B cell development.
A. Representative flow cytometry plots of the indicated surface markers in the bone marrow (BM) and spleen of *Zmynd8*$^{F/F}$ *Cd19*$^{Cre/+}$ and *Zmynd8*$^{+/+}$ *Cd19*$^{Cre/+}$ mice. **B.** Summary graphs for three independent experiments of the data shown in (B). **C.** WB analysis of WCE from un-activated (0h) mature *Zmynd8*$^{+/+}$ *Cd19*$^{Cre/+}$ B lymphocytes and after CSR induction in the presence of LPS and IL-4 for 48 and 72 h.

CSR of IgM to other isotypes can be induced in vitro by different cytokine treatments. Specifically, LPS and IL-4 stimulates switching to IgG1; LPS, TGFβ and BAFF promote IgG2b CSR; and LPS alone induces the expression of IgG3 in the surface of B cells (*see section 1.5.1).* WT and ZMYND8-deficient splenocytes were cultured in the presence of these different cytokine cocktail mixes to study the requirement of ZMYND8 for CSR in primary B cells. ZMYND8 deficiency lead to a significant and consistent reduction of all antibody isotypes tested *ex vivo* (**Figure 4.7A-C**).

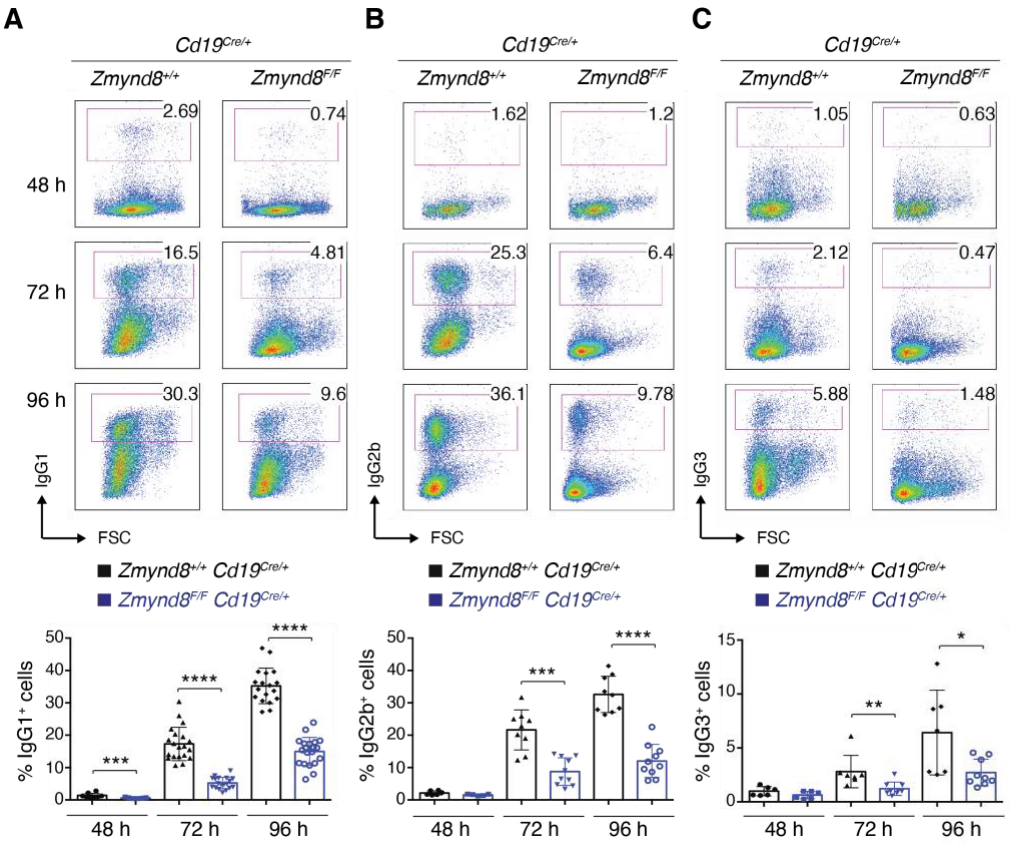

Figure 4.7. ZMYND8 is a novel factor required for CSR *ex vivo*.
A-C. Top: representative flow cytometry plots indicating the percentage of IgG1 (LPS and IL-4) (**A**), IgG2b (LPS, TGFβ and BAFF) (**B**) and IgG3 (LPS) (**C**) positive cells of *Zmynd8*$^{F/F}$ *Cd19*$^{Cre/+}$ and *Zmynd8*$^{+/+}$ *Cd19*$^{Cre/+}$ splenocytes stimulated for the indicated time points. Bottom: graph summarizing the data of at least five mice per time point and per genotype. Significance was calculated with the Mann-Whitney U test and error bars represent SD. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001.

To further understand if the CSR defect observed in cultured splenocytes is a reflect of a defective immune response *in vivo*, control and *Zmynd8* conditional KO mice were immunized. Antibody titers present in the serum 7 days (peak of germinal center formation) and 20 days post-immunization were determined by ELISA (enzyme-linked immunosorbent assay). Specifically, NP-CGG antigen (4-Hydroxy-3-nitrophenylacetyl hapten conjugated to chicken gamma globulin) was used to induce a T cell-dependent response that promotes B cell activation, GC formation and, therefore, SHM and CSR[249]. As shown in **Figure 4.8**, ZMYND8 deficiency lead to a significant decrease in NP-specific IgG1 antibodies compared to control, both at day 7 and day 20 after immunization. NP-specific IgM titers were slightly decreased at the immunization day (day 0) and merely increased at day 7, which correlates with the reduced ability of ZMYND8-deficient cells to undergo CSR. Altogether, these data evidenced that ZMYND8 is required for efficient CSR and the establishment of an effective immune response.

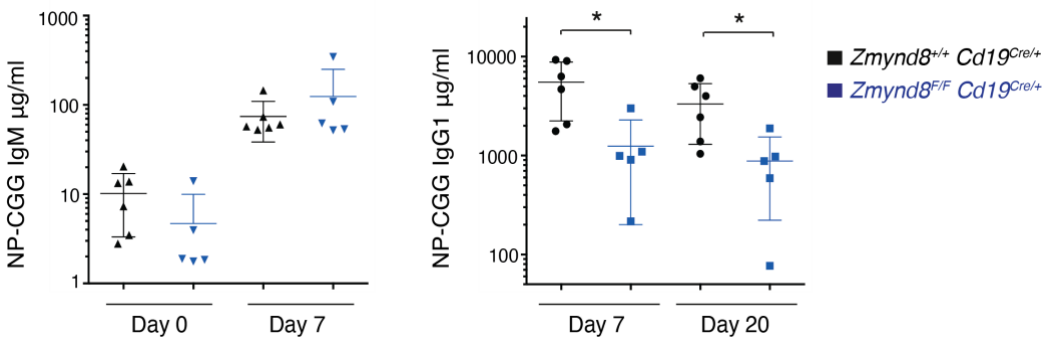

Figure 4.8. ZMYND8 is a novel factor required for CSR *in vivo*.
Summary plots for the titers of NP-reacting IgM (left) and IgG1 (right) antibodies in the serum of *Zmynd8*$^{F/F}$ *Cd19*$^{Cre/+}$ and *Zmynd8*$^{+/+}$ *Cd19*$^{Cre/+}$ before (Day 0) and after 7- and 20-days post-immunization with NP-CGG antigen. Graph summarizing the data of at least five mice per time point and per genotype. Significance was calculated with the Mann-Whitney U test and error bars represent SD. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001.

4.2.2 ZMYND8 is dispensable for B cell proliferation

B cell proliferation and cell cycle progression is a limiting step for CSR to happen, since undivided B-lymphocytes do not undergo antibody isotype differentiation and AID expression is cell-division dependent (*see section 1.5.5)*[93,107,111].To assess if ZMYND8 plays a role in cell cycle progression, WT and ZMYND8-deficient splenocytes were labeled with Cell Trace Violet (CTV). This fluorescent dye can permeate the plasma membrane and form covalent bonds to free amines that are present on the cellular surface and intracellularly. Upon cell division, the dye content gets diluted by half and changes can be tracked daily by flow cytometry, allowing the comparison of cells from different genotypes[250]. Splenocytes were isolated, labeled with CTV and stimulated with different cytokines to undergo CSR to IgG1, IgG2b and IgG3. CTV dilution profiles from ZMYND8-deficient B cells were comparable to those from WT at all time points measured (**Figure 4.9A-C**) and CSR was reduced in all cell cycle divisions upon ZMYND8 deletion (**Figure 4.9D**). Moreover, no significant changes in cell growth were observed upon CSR induction of splenocytes from the two genotypes (**Figure 4.9E**), leading to the conclusion that ZMYND8 role in antibody isotype differentiation is not related to B cell division.

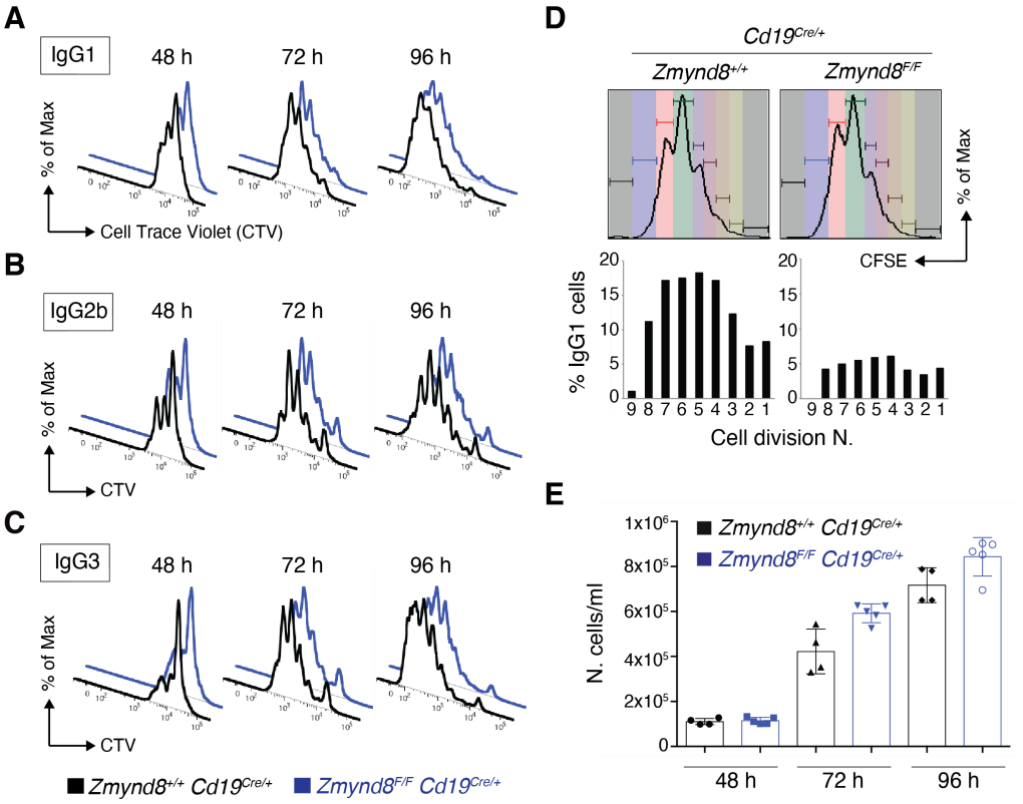

Figure 4.9. ZMYND8 is dispensable for B cell proliferation.
A-C. Representative histogram flow cytometry plots showing Cell Trace Violet (CTV) dilution of labeled splenocytes of the indicated genotypes stimulated for 48, 72 and 96 h in the presence of LPS and IL-4 (A), LPS, TGFβ and BAFF (B) and LPS (C). Data were derived from at least two mice per genotype. **D.** Flow cytometry analysis indicating the percentage of IgG1 positive cells in each cell division as measured by CFSE dye dilution. Primary B lymphocytes were isolated from *Zmynd8*$^{F/F}$ *Cd19*$^{Cre/+}$ and *Zmynd8*$^{+/+}$ *Cd19*$^{Cre/+}$ mice and cultured for 72h with LPS and IL-4. Data are representative of two independent experiments with at least four mice per genotype. **E.** Graph summarizing the number of cells per milliliter of splenocytes from the indicated genotypes cultured for 48, 72 and 96 h in the presence of LPS and IL-4. Data are representative of two independent experiments with at least four mice per genotype.

4.2.3 ZMYND8 does not regulate the repair of DSBs

Upon AID-mediated DSB formation within the S regions, the DDR is activated, and DNA repair factors mediate c-NHEJ, leading to productive CSR events. Alternatively, these DSBs can be internally re-joined preferentially with the use of microhomologies that are exposed by limited resection within the S regions, which are internally repetitive[94,95]. By protecting DNA ends against resection, phosphorylated 53BP1[94,95,98,101,147,148] and its downstream effector RIF1[96,97,100,137] counteract this unproductive micro-homology-mediated intra-switch region repair and favor classical end-joining of distal S breaks (*see section 1.5.4 and 1.7*). To determine the consequences of ZMYND8 ablation in DNA damage-induced signaling cascade and DNA repair, *Zmynd8*$^{-/-}$ immortalized mouse embryonic fibroblasts (iMEFs) were generated by CRISPR/Cas9 gene targeting (**Figure 4.10A-B**). Of note, a "random" clone (*Rc*), was generated in parallel as a control by using

a gRNA not present in the mouse genome. iMEFs constitute an ideal model system for performing assays that require adherent cells and for determining the recruitment of DNA repair factors to sites of ionizing radiation (IR)-induced DNA damage, that can be observed microscopically as foci (in the majority of cases). Although ZMYND8 localized into the nucleus, it did not seem to be a focus forming protein (**Figure 4.10A**).

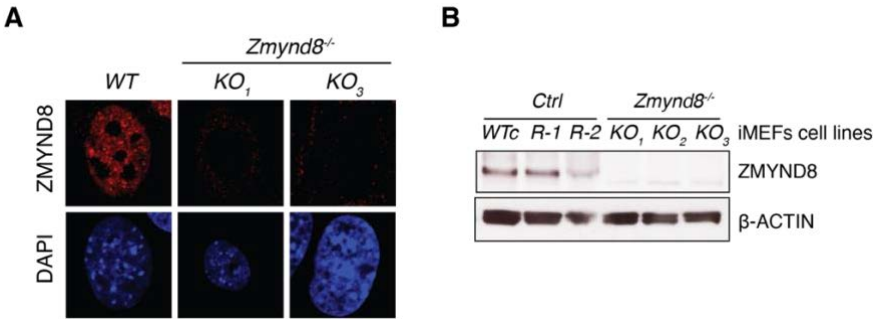

Figure 4.10. Generation of *Zmynd8*-KO iMEFs.
A-B. Representative immunofluorescent staining images (A) and western blot analysis (B) of *WT* and *Zmynd8*$^{-/-}$ immortalized mouse embryonic fibroblasts (iMEFs). Three different *Zmynd8*$^{-/-}$ clones derived from CRISPR/Cas9 somatic targeting (*KO1*, *KO2* and *KO3*) were compared to control cell lines generated in parallel: *WTc*, a WT clonal derivative; *R-1* and *R-2*, two independent WT clones nucleofected with a gRNA not present in the mouse genome.

Importantly, ZMYND8 deficiency did not affect the DNA-damage induced signaling cascade, as measured by immunofluorescence staining of H2AX phosphorylation (γH2AX) and 53BP1 and RIF1 foci formation (**Figure 4.11**).

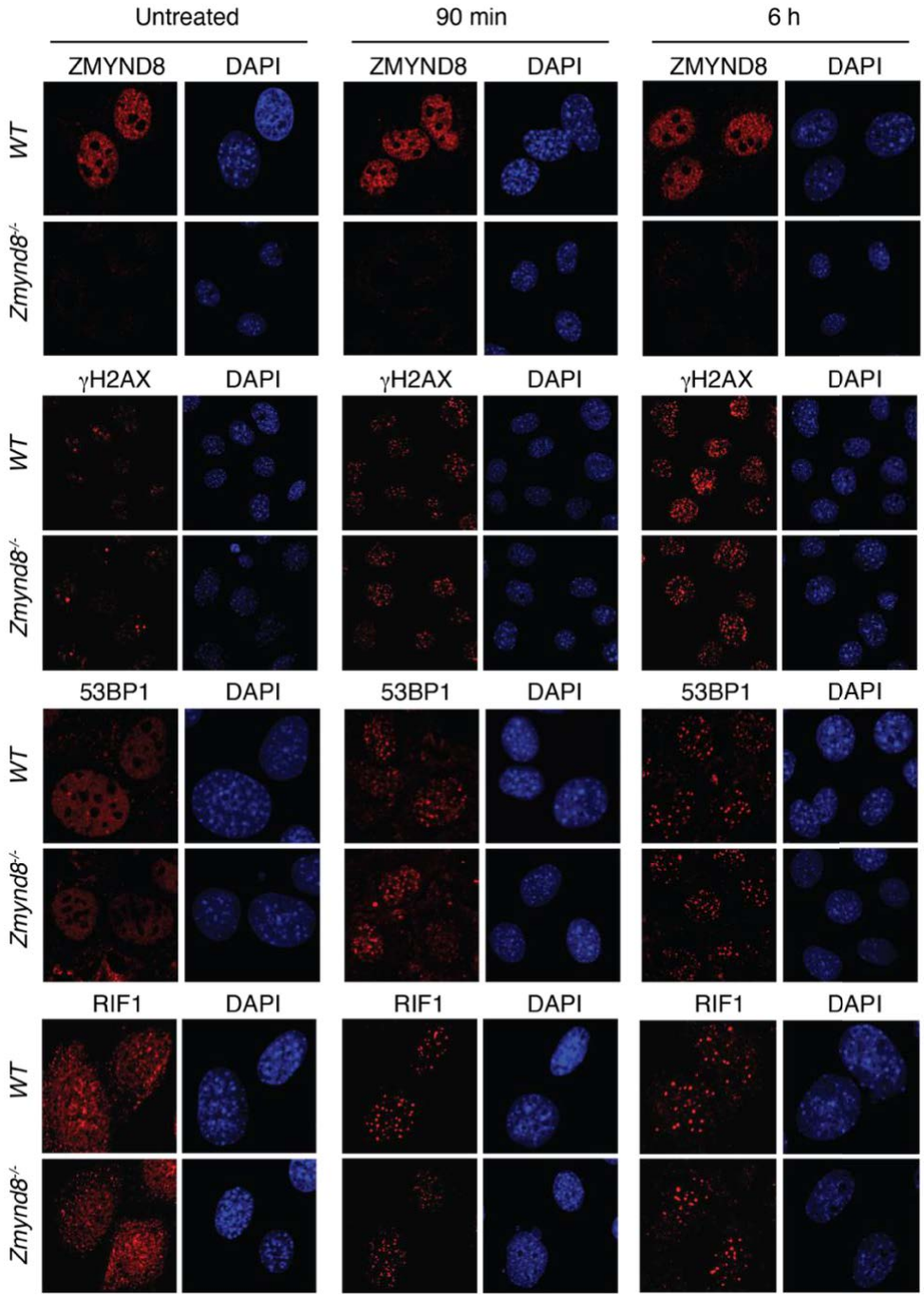

Figure 4.11. ZMYND8 is dispensable for the signaling of stochastic DNA damage.
Representative immunofluorescent staining images for the indicated factors of irradiation-induced foci (10 Gy) in *WT* and *Zmynd8*$^{-/-}$ iMEFs following 90 min and 6 h of recovery. Data derive from two independent experiments.

Since ZMYND8 had been previously suggested to play a role in DSB by HR[229–233], clonogenic survival assays following IR and PARP inhibition were performed. IR leads to breaks preferentially repaired by NHEJ, while PARPi induces DNA replication associated breaks that are physiologically repaired by HR[251,252]. ZMYND8-deficient iMEFs cell survival was similar to controls (*WT* and *Rc*) in both treatment scenarios (**Figure 4.12A-B**). In agreement with the lack of sensitivity to PARPi, *Zmynd8*$^{-/-}$ iMEFs

did not show any significant difference in chromosomal aberrations compared to ZMYND8-proficient cells in the presence of PARPi. *Brca1*$^{-/-}$ iMEFs were used as a positive control, since this gene is a major component of HR and KO cells die when treated with PARPi due to their inability to repair breaks that occur during DNA replication (**Figure 4.12C**)[253].

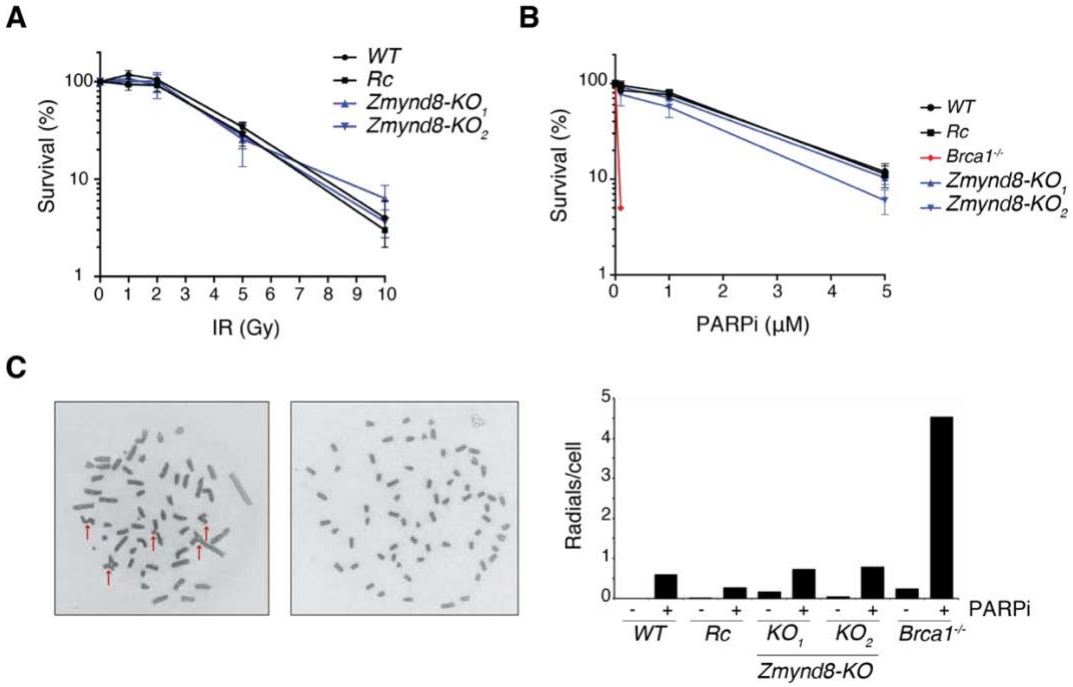

Figure 4.12. ZMYND8 is dispensable for the repair of IR- and PARPi-induced DSBs.
A-B. Graphs representing the percentage of cell survival as determined by colony formation assay of iMEFs cell lines of the indicated genotypes following ionized radiation (IR) (**A**) or PARP inhibition treatment (PARPi) (**B**). Each condition was performed in triplicates and error bars represent the mean. Data of two independent experiments per graph are shown. **C.** Representative metaphase images (left) and graph summarizing the number of radials per cell observed in iMEFs clones of the indicated genotypes in the presence or absence of 2 μM of PARPi for 24 h (right). Data of at least 42 metaphases per condition are shown. Red arrows indicate radials.

To determine if ZMYND8 is affecting the repair of DSBs in the context of antibody isotype differentiation, CSR was induced *via* CRISPR/Cas9 by targeting the Sμ (donor) and Sα (acceptor) regions in CH12 cells[254]. In this case, no cytokines were added to prevent the induction of AID-mediated DSBs. The percentage of IgA positive cells in ZMYND8-deficient cells was comparable to WT and also to *Aicda*$^{-/-}$ cells, which were used as a positive control. *Lig4*-, *Trp53bp1*- and *Rif1-KO* cells were unable to rescue CSR due to their role in c-NHEJ. In these cases, residual IgA levels could be explained by the activity of the alt-NHEJ pathway that can join the DNA ends by using microhomology regions when c-NHEJ is defective (**Figure 4.13**)[94,101,102]. Furthermore, the consequence of ZMYND8 ablation in the repair of AID-induced DSBs at the S regions was determined

by amplifying the Sμ-Sγ1 junctions of IgG1 switched primary B-lymphocytes. Specifically, the frequency of blunt-joined DNA ends together with nucleotides insertions or deletions and microhomology at the junction was analyzed. No significance differences were observed when comparing *Zmynd8*$^{F/F}$-*Cd19*$^{Cre/+}$ to *Cd19*$^{Cre/+}$ control splenocytes, indicating that ZMYND8 does not regulate microhomology mediated-end joining (MMEJ) of AID-induced breaks at the *Igh* locus (**Figure S1**). Altogether, these results demonstrated that ZMYND8 is dispensable for DSB-induced signaling and repair.

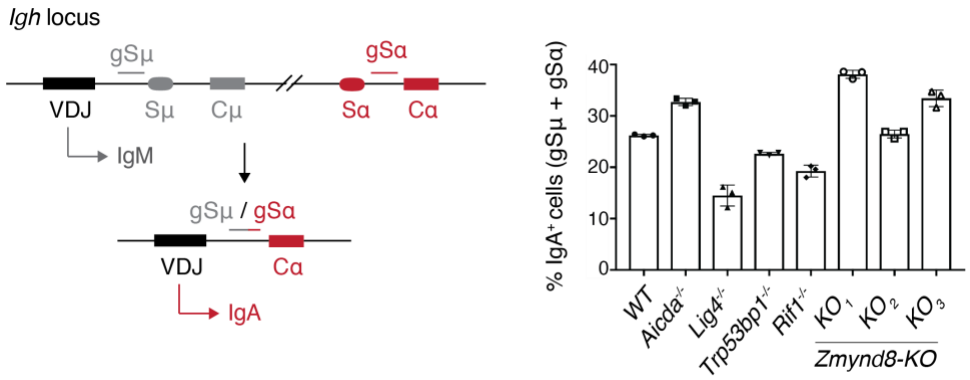

Figure 4.13. ZMYND8 is dispensable for the repair of Cas9-induced DSBs.
Left: experimental scheme of CSR assay induced via CRISPR/Cas9. Right: graph summarizing the percentage of IgA positive cells in un-activated CH12 cell lines of the indicated genotypes following somatic targeting of Switch-μ (gSμ) and Switch-α (gSα) regions. Data are representative of three independent experiments.

4.2.4 ZMYND8 is not essential for AID expression and function

In order for antibody isotype differentiation to efficiently occur, the two copies of the *Aicda* gene need to be expressed and the resulting AID protein has to target the S regions at the *Igh* locus. Therefore, *Aicda* haploinsufficiency and reduced AID expression or targeting result in CSR deficiency[255,256]. To first exclude the possibility that ZMYND8 is involved in *Aicda* transcription, *Aicda* mRNA levels in primary B cells were quantified by RT-qPCR (quantitative reverse transcription PCR). Differently to *Aicda-KO* cells that were used as a positive control, ZMYND8-deficient primary B-lymphocytes showed no apparent reduction in *Aicda* transcript levels compared to control (**Figure 4.14A**). To determine if ZMYND8 is regulating AID-dependent formation of DSBs, mutation rates at the 5'Sμ region were measured by paired-end next generation sequencing (MutPE-seq)[93,257]. This assay allows for the detection of mutations introduced by AID. Since the S regions are highly repetitive, deep sequencing was carried at the 5' extreme of Sμ, where a "footprint" of AID-mutations can be detected (**Figure 4.14B, left**). Moreover, amplification with low number of PCR cycles, the use of a proof-reading polymerase and the fact of having the information about the two strands of the DNA by pair-end

sequencing ensured the exclusive detection of AID-dependent introduced mutations. This assay was performed in splenocytes cultured in the presence of LPS and IL-4 for 72 or 96 h and no differences were observed in any of the cases when comparing ZMYND8-deficient and -proficient cells (**Figure 4.14B, right**). Altogether, these data support the conclusion that ZMYND8 is not required for efficient *Aicda* gene expression and that ZMYND8 deficiency does not regulate AID function.

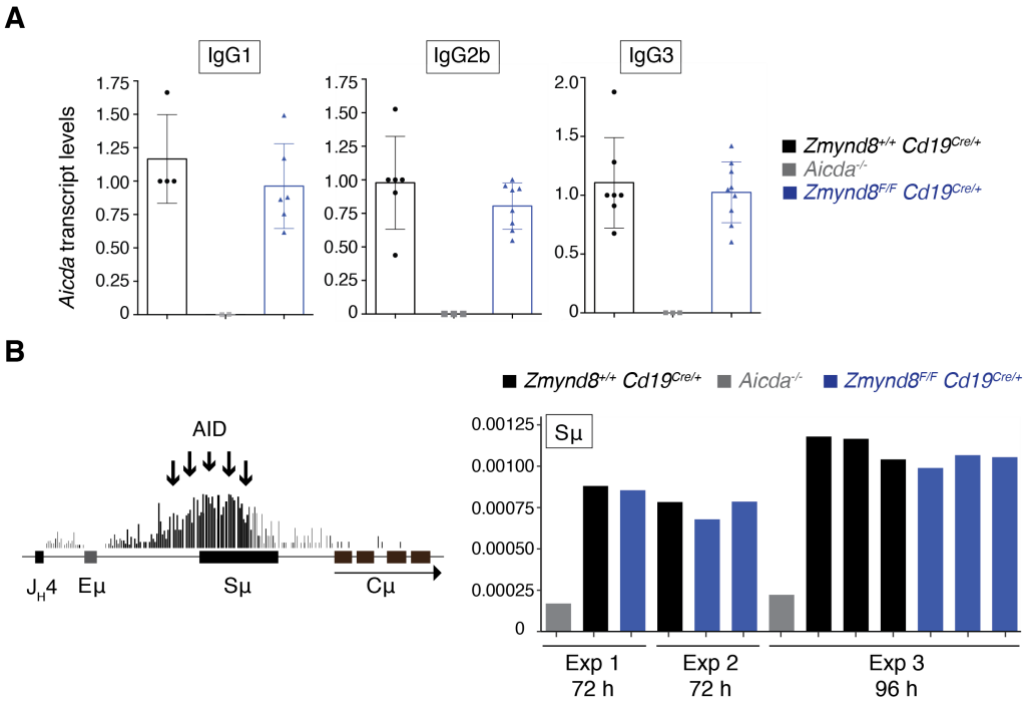

Figure 4.14. ZMYND8 is dispensable for AID expression and induction of DSBs.
A. RT-qPCR analysis of *Aicda* mRNA levels in splenocytes of the indicated genotypes stimulated to undergo CSR to IgG1 (LPS and IL-4), IgG2b (LPS, TGFβ and BAFF) and IgG3 (LPS) for 48 h. Data were normalized to a *Cd19*$^{Cre/+}$ control within each experiment and error bars indicate SD. **B.** Left: scheme representing the mutations generated by AID in the region of the *Igh* locus confined in between J_H4 and Cμ. Right: graph showing the relative mutation rates measured by MutPE-seq at 5'Sμ region in splenocytes of the indicated genotypes cultured in the presence of LPS and IL-4 for 72 or 96 h. Three independent experiments were performed, each column representing one mouse.

4.2.5 ZMYND8 promotes efficient transcription of acceptor S regions

AID targeting to the S regions depends on the noncoding transcription across the donor (Sμ) and acceptor S elements and the following constant genes (GLT)[51,52,62]. Sμ region is constitutively transcribed in naïve B cells and increases upon stimulation. On the other hand, GLT across the acceptor region(s) is induced in a cytokine dependent manner due to the accumulation of histone modifications that determine an open chromatin status (H3K4me3 and H3K9ac) at the same time the repressive marks H3K9me3 and H3K27me3 are removed[10,82–84]. Moreover, splicing of primary GLTs has been shown to

be required for efficient AID targeting to the S regions, although there are different models and still controversy in literature understanding the reason for it (*see section 1.5.3*)[58]. To determine if ZMYND8 deficiency is regulating GLT or playing a role in splicing, Sμ and pre- and post-spliced Igγ1, Igγ2b and Igγ3 transcripts were amplified by RT-qPCR analysis on splenocytes stimulated to undergo CSR to IgG1, IgG2b and IgG3, respectively (**Figure 4.15**). Of note, cells were collected 48 h post-activation because GLT is an early event during antibody isotype differentiation.

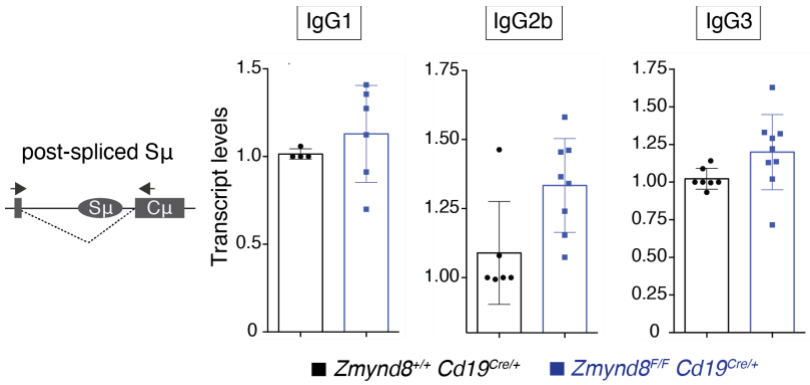

Figure 4.15. ZMYND8 is dispensable for Sμ (donor) GLT.
Left: schematic representation of the position of the primers used to amplify post-spliced Sμ GLTs. Right: graphs representing the GLT levels of post-spliced Sμ in primary B-lymphocytes stimulated with LPS and IL-4 (IgG1), LPS, TGFβ and BAFF (IgG2b) and LPS (IgG3) for 48 h. Data were normalized to a *Cd19*Cre/+ control within each experiment.

Whereas Sμ GLT in *Zmynd8-KO* cells was undistinguishable from control, pre- and post-spliced acceptor GLT were significantly reduced in all isotypes tested upon *Zmynd8* deletion (**Figure 4.16A-B**). Conclusively, ZMYND8 is required for efficient transcription of acceptor S regions and, therefore, AID targeting at the *Igh* locus.

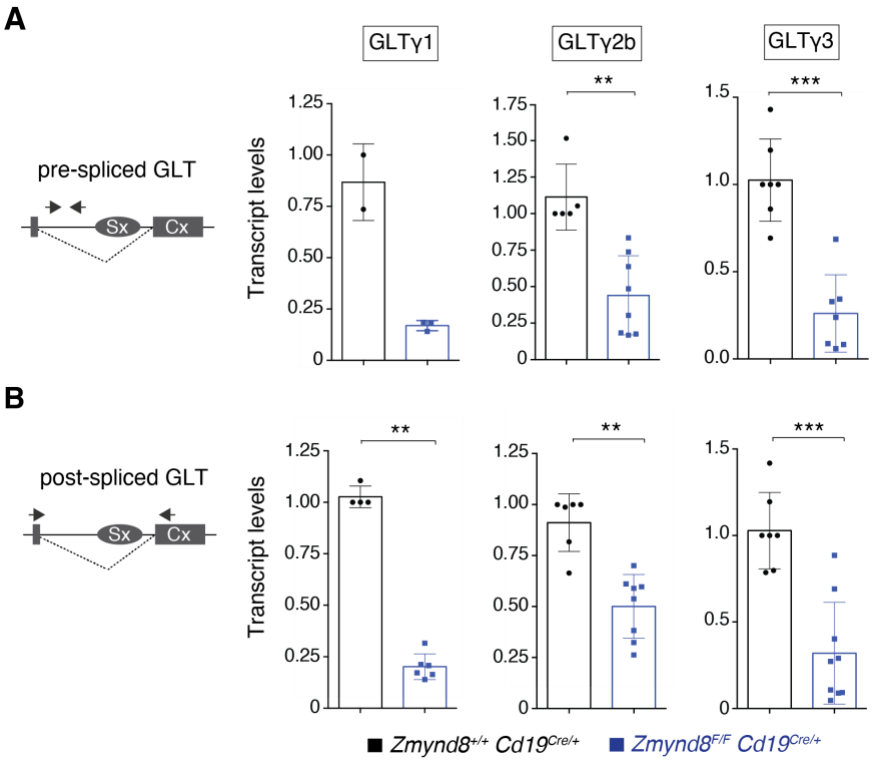

Figure 4.16. ZMYND8 supports GLT of acceptor S regions.
Left: schematic representation of the position of the primers used to amplify pre- and post-spliced acceptor GLTs. Right: graphs representing the GLT levels of pre and post-spliced Sγ1, Sγ2b and Sγ3 in primary B lymphocytes stimulated with LPS and IL-4 (IgG1), LPS, TGFβ and BAFF (IgG2b) and LPS (IgG3) for 48 h. Data were normalized to a *Cd19*$^{Cre/+}$ control within each experiment. Significance was calculated with the Mann-Whitney U test and error bars represent SD. **p ≤ 0.01, ***p ≤ 0.001.

4.2.6 ZMYND8 associates to the 3' *Igh* enhancer and regulates its transcriptional activity

ZMYND8 is a chromatin reader that associates with enhancers and super-enhancers through its interaction with H3K4me0, H3K4me1, H3K14Ac and H3K27Ac histone marks, promoting transcriptional repression and, in some cases, activation[225,227–230,232]. To further dissect the role of ZMYND8 in CSR, its association with the *Igh* locus was determined by chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq). Consistent with previous literature findings, ZMYND8 binding to the chromatin was enriched in the two major *Igh* super-enhancer regions: Eμ, located 5' of Sμ; the uncharacterized enhancer in between Cγ1 and Sγ2b; and the 3' regulatory region (3'RR), found at the 3' end of the *Igh* locus and composed of four transcriptional enhancers (DNase 1 hypersensitive sites hs3a, hs1,2, hs3b and hs4) (*see section 1.6*) (**Fig. 13A**)[17–19,25,27]. Since ZMYND8 binding to chromatin was reported to have an effect on transcription, RNA polymerase II (RNAPII) loading at the *Igh* locus was measured by ChIP-seq analysis in stimulated ZMYND8-deficient and -proficient B cells. Importantly, the C-terminal domain of RNAPII is phosphorylated at different sites to modulate its

activity during transcription. Consequently, phosphorylation at Serine5 (RNAPII-S5) is enriched at promoters and transcription start sites of genes, signifying that the polymerase has been loaded. As transcription elongation takes place, S5 phosphorylation is progressively removed at the same time S2 is phosphorylated[258]. Specifically, ZMYND8-deficient primary B-lymphocytes and CH12 cells showed a significant increase in RNAPII-S5 at the 3'RR compared to control, being the effect more prominent in hs1,2 and hs3b core enhancer regions (**Figure 4.17A-B** and **Figure S3**).

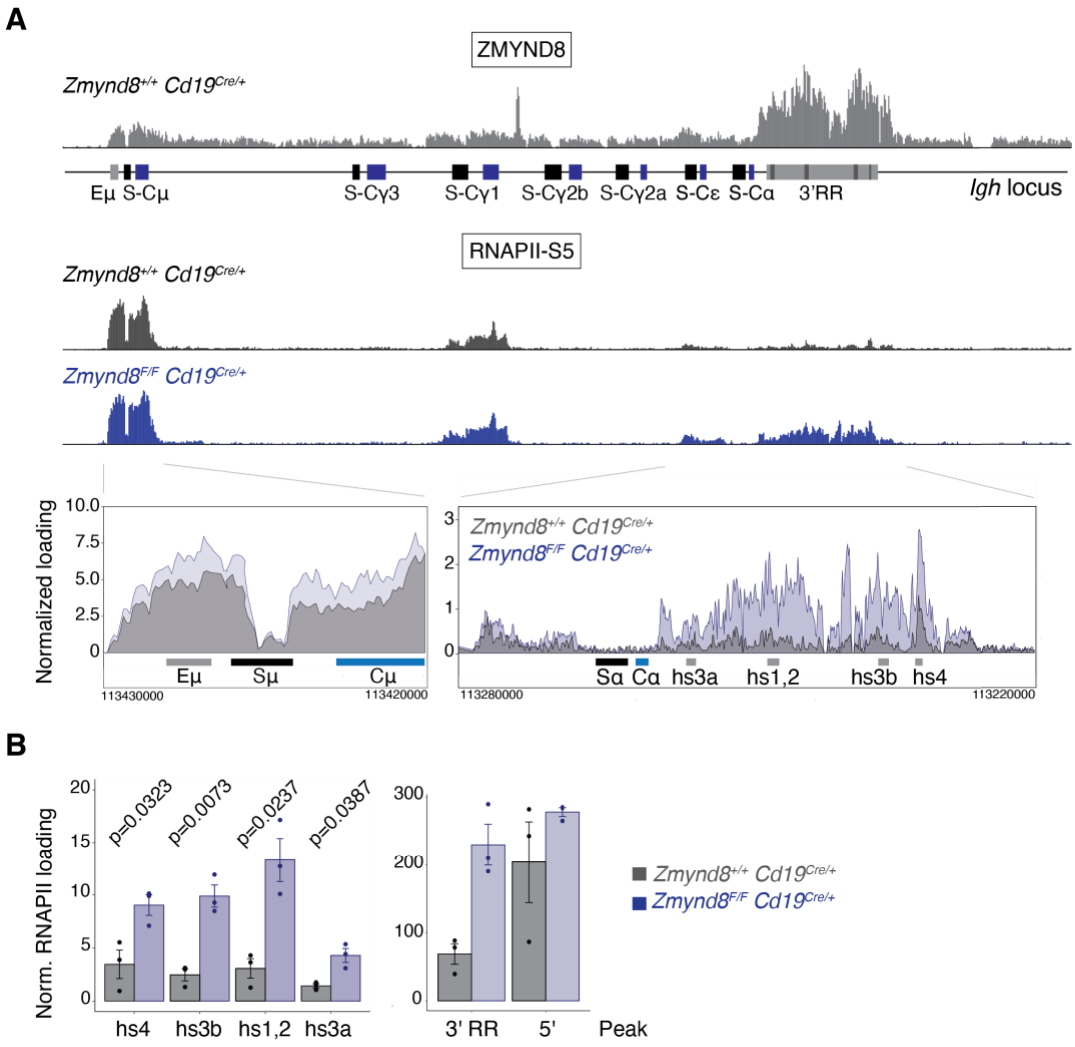

Figure 4.17. ZMYND8 binds to *Igh* enhancers and regulates RNAPII loading at the 3'RR. **A.** Top: ZMYND8 and RNA polymerase II phosphorylated on serine 5 (RNAPII-S5) loading at the Igh locus as measured by ChIP-seq analysis in naïve B cells stimulated to undergo CSR to IgG1 for 72 h. Bottom: overlaid RNAPII-S5 ChIP-seq tracks of *Zmynd8$^{F/F}$ Cd19$^{Cre/+}$* and *Zmynd8$^{+/+}$Cd19$^{Cre/+}$* mice at Eμ and 3'RR enhancer regions. **B.** Quantification of RNAPII-S5 occupancy at the 3'RR core enhancer regions hs4, hs3b, hs1,2 and hs3a (left) and at Eμ and 3'RR enhancers (right). Data derived from 3 independent mice per genotype and error bars indicate SD. Significance was calculated with Welch 2 sample unpaired t-test.

RT-qPCR analysis of hs1,2 and hs3b in *Zmynd8*$^{F/F}$-*Cd19*$^{Cre/+}$ and *Cd19*$^{Cre/+}$ splenocytes stimulated to undergo CSR to IgG1 and IgG3 confirmed that the increase in RNAPII loading at the 3'RR upon *Zmynd8* deletion corresponded to increased transcription levels of eRNAs (enhancer RNAs) (**Figure 4.18**). Furthermore, RNA-seq analysis of two independent *Zmynd8-KO* CH12 clones provided the same results (**Figure S2**). Therefore, ZMYND8 binds to the 3'RR and modulates its transcriptional status by preventing its overexpression.

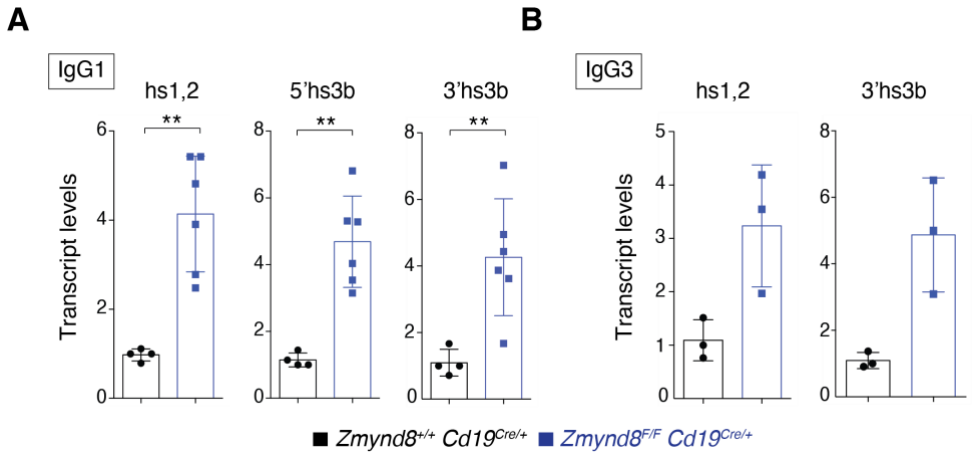

Figure 4.18. ZMYND8 represses hs1,2 and hs3b transcription.
A-B. RT-qPCR analysis of hs1,2 and hs3b eRNA levels in splenocytes stimulated to undergo CSR to IgG1 (A) or to IgG3 (B). Two different primer pairs located at the 5' (5'hs3b) and at the 3' (3'hs3b) regions of the hs3b core enhancer were used. Data were normalized to a *Cd19*$^{Cre/+}$ control within each experiment. Significance was calculated with the Mann-Whitney U test and error bars represent SD. **p ≤ 0.01.

4.2.7 ZMYND8 regulates SHM of the heavy chain locus

The 3'RR has been shown to be required for efficient GLT as well as for SHM of the heavy chain locus (*see section 1.6*)[22,24–27,128,129,134]. SHM is the responsible process for increasing antigen affinity of antibodies and it is also dependent on AID, which introduces mutations in the variable regions of the heavy and light chain locus (*see section 1.4*)[55,56]. To further characterize the role of ZMYND8 in the 3'RR activity, the number of AID-induced mutations in the variable regions of heavy and light chains were quantified. This experiment was performed in aged mice because Peyer's Patches (accumulations of lymphoid follicles located in the intestine) contain activated B-lymphocytes that undergo SHM due to their continue exposure to food and exogenous antigens[259–261]. Specifically, this process happens in the so-called germinal centers (GCs), which are microanatomic structures of secondary lymphoid organs where SHM and CSR take place upon antigen encountering[43]. There are two possible light chains that can be rearranged and form a B cell receptor (BCR) together with the heavy chain during B cell development, Igκ and

Igλ. While human B cells present a ratio of 1.5:1 of Igκ to Igλ, murine B cells have a ratio of 19:1, which explains why the Igκ is considered over the Igλ when studying SHM in mice (*see section 1.2.2*)[32]. Genomic DNA was extracted from sorted GC and non-GC (NGC) B cells and the intronic regions located immediately downstream the recombined V(D)J in the heavy (*Igh*) and in the light (*Igκ*) loci were amplified and sequenced[259,262,263]. Amplicons were aligned to un-mutated J_H4 and $J_κ5$ sequences and the number of point mutations were quantified. As expected, only GC B cells contained mutations, as SHM does not take place outside the GC (**Figure 4.19A and 4.19D**). Surprisingly, ZMYND8-deficient GC B cells showed a decreased number of mutated sequences and mutation frequency as compared to $Cd19^{Cre/+}$ cells at the 3'J_H4 region (**Figure 4.19A-B and 4.19E**). Moreover, GC B cells presented a bias towards low-mutated sequences and a significant decrease in high rate of mutations in this region upon *Zmynd8* deletion. However, similar number and distribution of mutations were found in $Zmynd8^{F/F}$-$Cd19^{Cre/+}$ and $Cd19^{Cre/+}$ GC B lymphocytes at the 3'J_K5 region, which is not regulated by the 3'RR (**Figure 4.19D and 4.19E**). Altogether, these data indicated that ZMYND8 is exclusively required for efficient SHM at the heavy chain locus, supporting the hypothesis that this factor regulates the activity and function of the 3'RR.

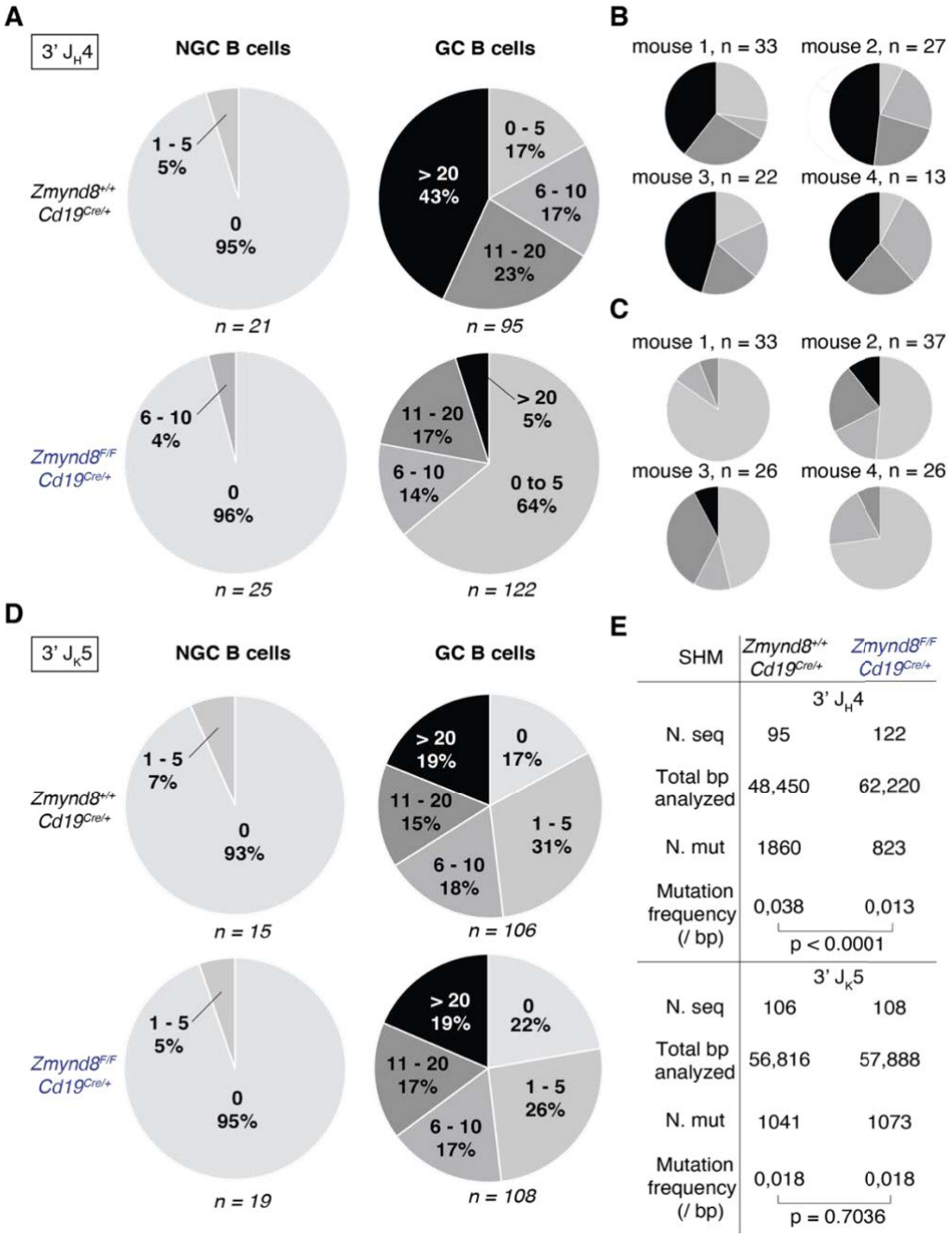

Figure 4.19. ZMYND8 is required for efficient SHM at the *Igh* locus.
Genomic DNA from non-germinal center (NGC) and germinal center (GC) B cells from the Peyer's Patches of un-immunized aged *Zmynd8*$^{F/F}$ *Cd19*$^{Cre/+}$ and *Cd19*$^{Cre/+}$ mice was amplified, cloned and analyzed (four mice per genotype). **A.** Cumulative pie charts representing the percentage of mutations in the 3'J_H4 intronic region over 510bp sequence. **B-C.** Pie charts showing the percentage of mutations in the 3'J_H4 intronic region of GC B cells from 4 independent *Cd19*$^{Cre/+}$ (B) and aged *Zmynd8*$^{F/F}$ *Cd19*$^{Cre/+}$ (C) mice. **D.** Cumulative pie charts representing the percentage of mutations in the 3'J_K5 intronic region over 510bp sequence. **E.** Table specifying the number of sequences (N. seq), total bp analyzed, number of mutations (N. mut) and mutation frequency (per bp) in the 3'J_H4 and 3'J_K5 intronic regions of GC B cells shown in (A) and (D). Significance was calculated with the Mann-Whitney U test.

4.3 PDAP1 is a novel CSR factor required for efficient AID expression and suppression of the integrated stress response

4.3.1. PDAP1 required for CSR

Since PDAP1 resulted a candidate of interest from the loss-of-CSR screen (**Figure 4.3B**), *Pdap1*$^{-/-}$ CH12 clones were generated to validate the observed phenotype and further characterize the role of this factor in antibody isotype differentiation. To do so, three independent paired gRNAs were designed and cloned into a plasmid containing the nickase version of Cas9 (Cas9^{D10A}) (**Figure 4.20A**). As expected, genomic *Pdap1* targeting led to a significant reduction of CSR (**Figure 4.20B**). Three clonal derivatives that showed a decrease in IgA levels were picked and their genomic scar and PDAP1 protein expression were analyzed. Two of these clones, P1 and P2, showed frameshift and premature termination codons in both alleles, resulting in lack of PDAP1 protein. P3 presented a 27 bp internal deletion within exon 4 that caused a decreased but still detectable PDAP1 protein (**Figure 4.20C**).

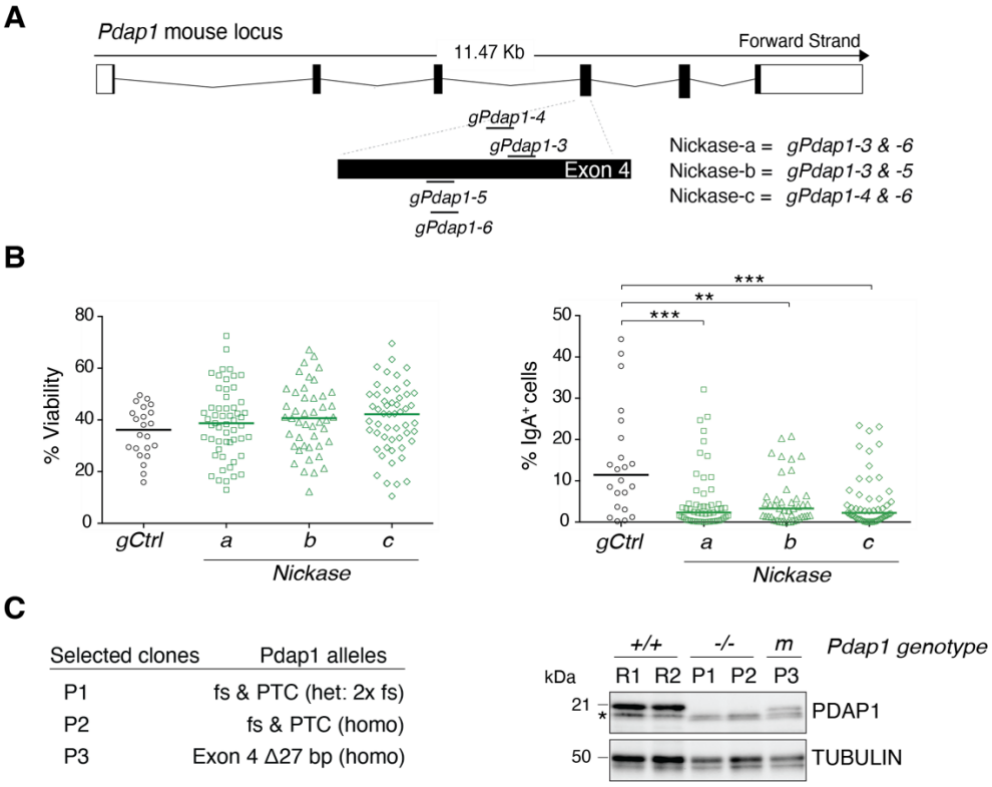

Figure 4.20. Generation of PDAP1-deficient CH12 clones.
A. Schematic representation of the murine *Pdap1* locus. The location of 3 different gRNA pairs used in this study that guide a Nickase-Cas9^{D10A} is indicated. **B.** Graphs showing the percentage IgA positive cells after somatic targeting of CH12 with g*Ctrl* (paired random gRNAs) and three independent gRNAs pairs targeting the Pdap1 locus (Nickase-a, -b and -c). Each dot indicates a single CH12 clone. **C.** Left: table listing the genomic conformation of 2 independent *Pdap1*$^{-/-}$ (P1, P2 and P2) and 1 mutant *Pdap1* (P3: *Pdap1*^{m}) CH12 clones selected from analysis shown in (B). PCT: premature termination codon; fs: frameshift; het: heterozygous; homo: homozygous; Δ:

deletion. Right: WB analysis of WCE from two random (R1 and R2) and the 3 Pdap1 clones described on top. Asterisk marks an unspecific band. Significance was calculated with the Mann-Whitney U test and error bars represent SD. **p ≤ 0.01, ***p ≤ 0.001.

Moreover, these CH12 clones exhibited a significant reduction in CSR compared to controls when activated *in vitro* (**Figure 4.21**), indicating that the deletion observed in P3 is not responsible for the loss of CSR caused by PDAP1. Altogether, these data confirm that PDAP1 is required for efficient antibody isotype differentiation in CH12 cells.

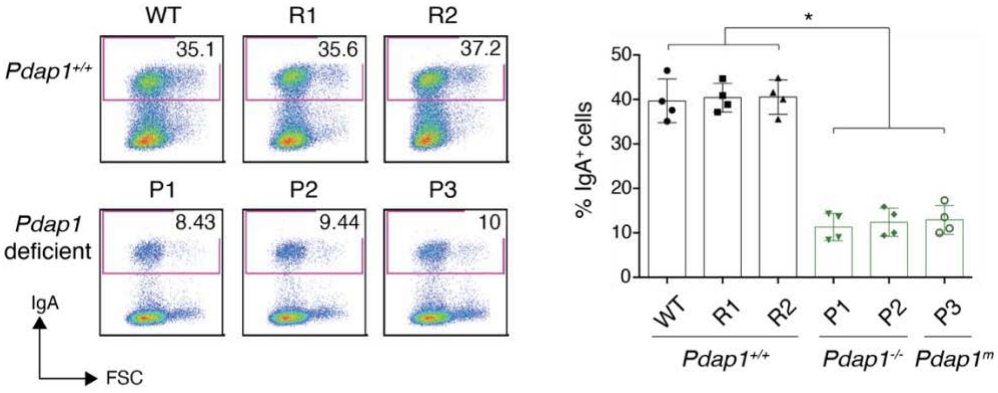

Figure 4.21. PDAP1 is a novel factor required for CSR in CH12.
Left: Representative flow cytometry plots showing the percentage of IgA+ cells from WT (parental cell line, R1 and R2), *Pdap1*$^{-/-}$ and *Pdap1*^{m} CH12 cells following 48 h of stimulation with αCD40, TGFβ and IL-4. Right: Graph summarizing 4 independent experiments. Significance was calculated with the Mann-Whitney U test and error bars represent SD. *p ≤ 0.05.

To validate the results obtained in CH12 and study the mechanistic aspects of PDAP1 in CSR, a conditional *Pdap1*-KO mouse model was generated. Specifically, two LoxP sites flaking the exon 2 were introduced by CRISPR-Cas9 targeting and subsequent repair by HR. Importantly, the targeting vector contained silent mutations at the gRNA sequence site in order for Cas9 not to cut it (represented with an asterisk in the figure) (**Figure 4.22A**). The resulted *Pdap1*$^{F/+}$ mouse was bred to get a homozygous *Pdap1*$^{F/F}$ allele, which was then crossed to a *Cd19*$^{Cre/+}$ mouse to conditionally delete *Pdap1* in B-lymphocytes at early stages of B cell differentiation[188]. The conditional deletion of *Pdap1* and the lack of protein levels were assessed by PCR and WB analysis, respectively (**Figure 4.22B-C**).

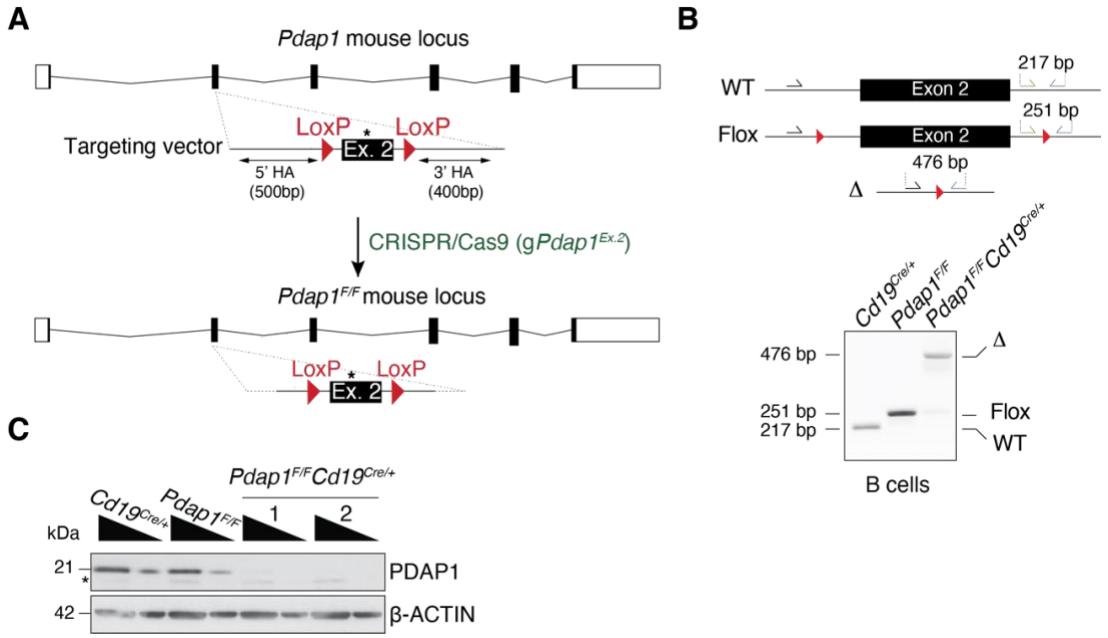

Figure 4.22. Generation of a conditional *Pdap1*-KO mouse model.
A. Representation of the murine *Pdap1* locus before and after CRISPR/Cas9 somatic targeting. Asterisk indicates silent mutations introduced within Cas9 target sites to exclude cutting of the targeting vector. HA: homology harm; LoxP: lox site. **B.** Top: schematic representation of the WT, Flox and deleted (Δ) *Pdap1* alleles and the location of primers to detect these alleles. Bottom: PCR analysis of gDNA extracted from splenocytes of the specified genotypes and amplified with the primers indicated on (top). **C.** WB analysis of WCE from splenocytes of the indicated genotypes cultured for 96 h in the presence of LPS and IL-4. Triangles represent threefold dilution. Asterisk indicates an unspecific band.

To understand if PDAP1 has any role in B cell development, surface expression of different markers that are expressed during B cell maturation in the bone marrow and spleen were determined by flow cytometry. PDAP1-deficient B-lymphocytes differentiated comparably to control, indicating that this factor does not play a role in V(D)J recombination or early B cell maturation (**Figure 4.23A-B**). However, a significant decreased number of mature B-lymphocytes ($CD43^-$) was observed in the absence of PDAP1 (**Figure 4.23C**). This can be due to alterations in late B cell differentiation or to apoptosis.

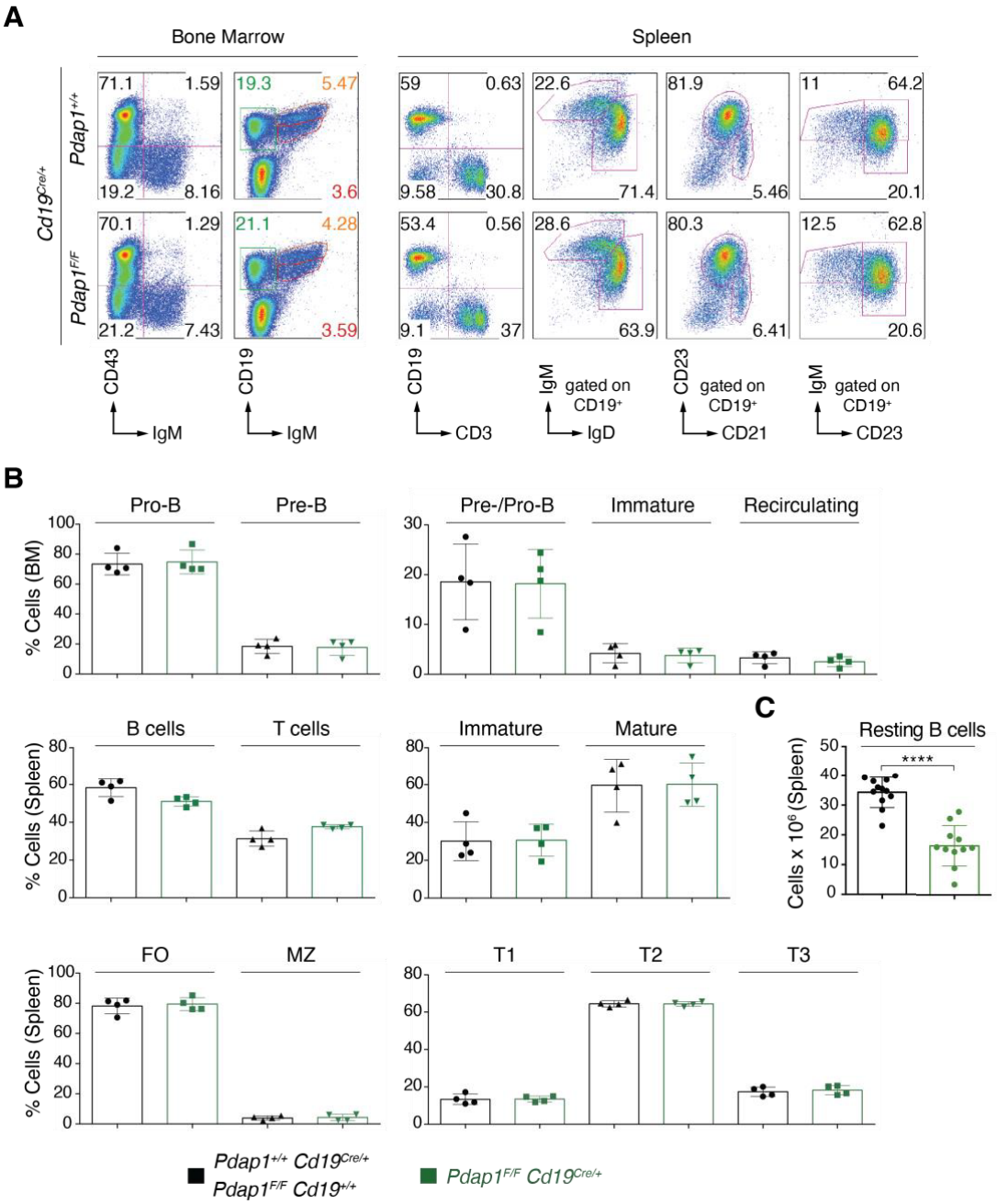

Figure 4.23. PDAP1 is largely dispensable for B cell development.
A. Representative flow cytometry plots of the indicated surface markers in the bone marrow (BM) and spleen of *Pdap1*$^{F/F}$ *Cd19*$^{Cre/+}$ and *Pdap1*$^{+/+}$ *Cd19*$^{Cre/+}$ mice. *Pdap1*$^{F/F}$ *Cd19*$^{+/+}$ were also used as a control. **B.** Summary graphs for 2 independent experiments of the data shown in (A). **C.** Number of resting B (CD43-) cells isolated from spleens of the genotypes indicated in (B). Each dot represents a different mouse. Significance was calculated with the Mann-Whitney U test and error bars represent SD. ***p ≤ 0.001.

To study the requirement of PDAP1 for CSR in primary B cells, WT and PDAP1-deficient splenocytes were cultured in the presence of different cytokine cocktail mixes and IgG1, IgG3, IgG2b and IgA levels were measured over time. The resulted data indicated that PDAP1 is necessary for efficient CSR to all antibody isotypes tested *ex vivo* (**Figure 4.24A-D**).

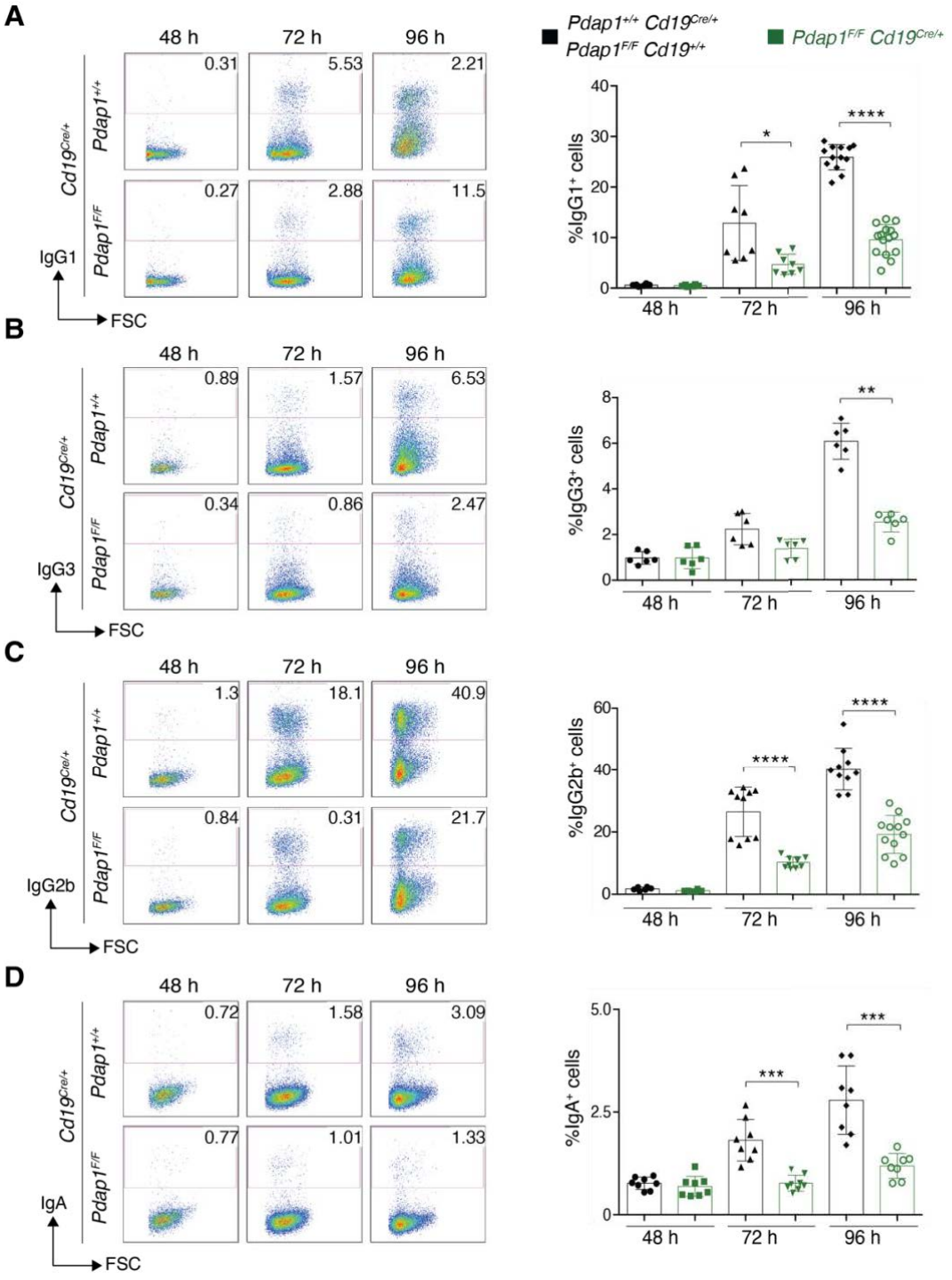

Figure 4.24. PDAP1 is required for efficient CSR *ex vivo.*

A-D. Left: Representative flow cytometry plots indicating the percentage of IgG1 (LPS and IL-4) (**A**), IgG3 (LPS) (**B**), IgG2b (LPS, BAFF, TGFβ) (**C**) and IgA (LPS, BAFF, TGFβ, IL-5) (**D**) positive cells of *Pdap1*$^{F/F}$ *Cd19*$^{Cre/+}$, *Pdap1*$^{+/+}$ *Cd19*$^{Cre/+}$ and *Pdap1*$^{F/F}$ *Cd19*$^{+/+}$ splenocytes stimulated for the indicated time points. Right: Graph summarizing the data of at least six mice per time point and per genotype. Significance was calculated with the Mann-Whitney U test and error bars represent SD. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001.

To further understand if the CSR defect shown in PDAP1-deficient cultured splenocytes was a reflect of a defective immune response *in vivo*, the percentage of IgA positive cells

was determined in the Peyer's Patches of control and PDAP1-deficient B-lymphocytes. As a result of continuous exposure to food, Peyer's Patches (located in the intestine) exhibit GC formation, thus containing B cells that undergo SHM and CSR to IgA[261]. In agreement with the data obtained *ex vivo*, B-lymphocytes displayed a significant reduction in IgA in the absence of PDAP1 as compared to control. Surprisingly, PDAP1-deficient mice also showed a decrease in the percentage of GC B cells. Conclusively, these data evidenced that PDAP1 is required for effective CSR *in vivo* (**Figure 4.25**).

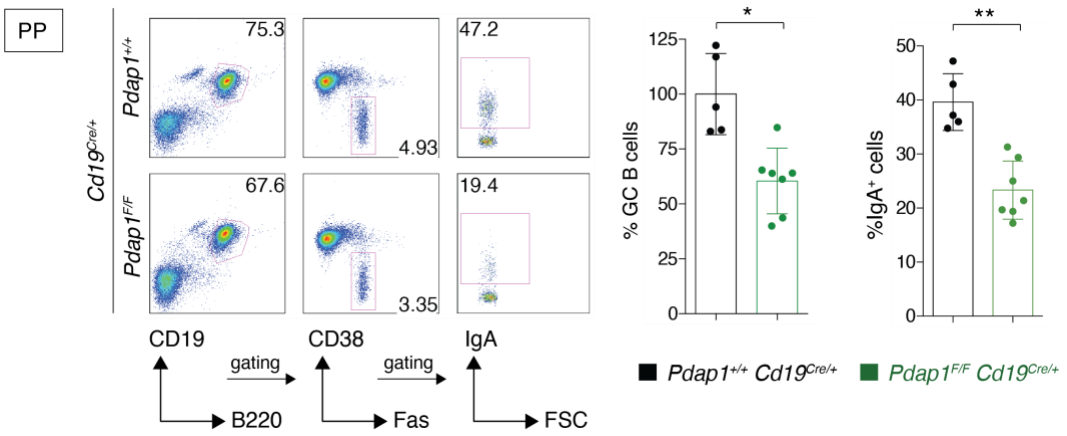

Figure 4.25. PDAP1 is required for efficient CSR *in vivo.*
A-D. Left: Representative flow cytometry plots showing the percentage of germinal center (GC) B cells ($CD19^+$ $B220^+$ $CD38^-$ Fas^+) and IgA positive splenocytes in Peyer's Patches (PP) of unimmunized mice. Right: Graph summarizing the data of at least five mice per genotype in two independent experiments. Significance was calculated with the Mann-Whitney U test and error bars represent SD. *p ≤ 0.05, **p ≤ 0.01.

4.3.2. PDAP1 is largely dispensable for B cell proliferation

To determine if the CSR reduction observed in PDAP1-deficient cells is due to a defect in B cell proliferation, splenocytes were labeled with Cell Trace Violet (CTV) and cultured in the presence of different cytokine mixes. The CTV dilution profiles of $Pdap1^{F/F}Cd19^{Cre/+}$ B cells were comparable to those from WT at all time points measured under all stimulation conditions (**Figure 4.26A-D**). Furthermore, CSR was reduced in all cell cycle divisions upon PDAP1 deletion (**Figure 4.26E**). Therefore, these data proved that PDAP1 role in CSR is not related to B cell proliferation.

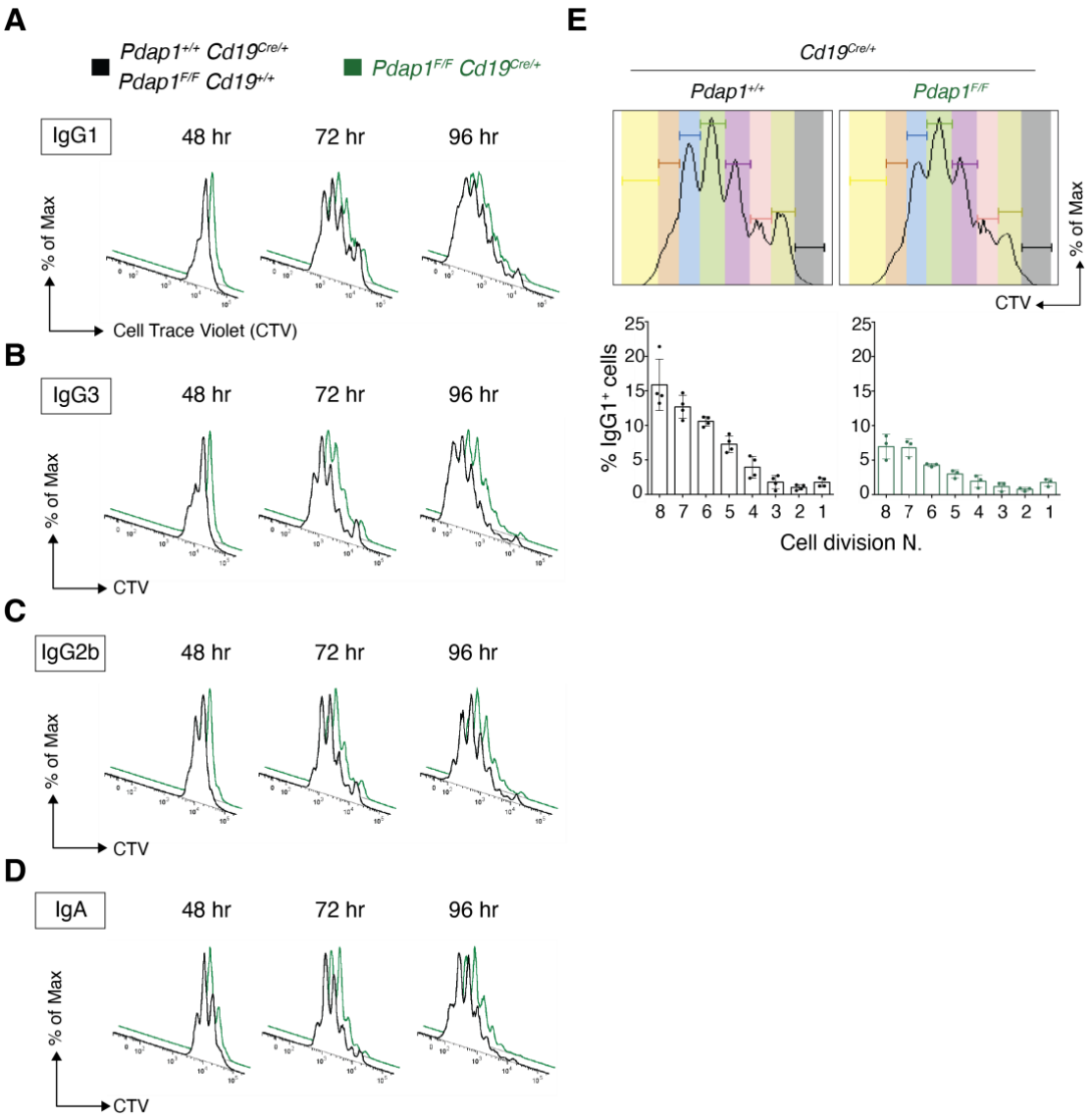

Figure 4.26. PDAP1 is dispensable for B cell proliferation.
A-D. Representative histogram flow cytometry plots showing Cell Trace Violet (CTV) dilution of labeled splenocytes of the indicated genotypes stimulated for 48, 72 and 96 h in the presence of LPS and IL-4 (**A**), LPS, (**B**) LPS, TGFβ and BAFF (**C**) and LPS, TGFβ, BAFF and IL-5 (**D**). Data were derived from at least two mice per genotype. **E.** Flow cytometry analysis indicating the percentage of IgG1 positive cells in each cell division as measured by CTV dye dilution. Primary B-lymphocytes were isolated from mice of the indicated genotypes and cultured for 72 h with LPS and IL-4. Data are representative of two independent experiments with at least three mice per genotype.

4.3.3. PDAP1 is required for efficient AID expression

Transcription across the donor (Sμ) and acceptor S regions together with the transcription and expression of the two *Aicda* copies are a requisite for productive CSR[70,255,256]. To determine if PDAP1 regulates any of these processes, GLT and *Aicda* levels were measured by RT-qPCR analysis in splenocytes activated *in vitro* to undergo CSR to different isotypes. Whereas GLT of the donor Sμ region was unaffected, *Pdap1*$^{F/F}$-*Cd19*$^{Cre/+}$ B cells exhibited a differential post-spliced germline transcript

expression depending on the switching conditions: Igγ1 and Igα GLT were significantly reduced in the absence of PDAP1; Igγ3 transcript was also decreased, although the effect was not as pronounced as in Igγ1 and Igα; and, lastly, Igγ2b GLT levels were comparable to control cells (**Figure 4.27A-B**). Moreover, a consistent two-fold reduction of *Aicda* mRNA levels were observed in PDAP1-deficient splenocytes switching to all isotypes tested (IgG1, IgG3, IgG2b and IgA) (**Figure 4.27C**). Notably, this decrease in AID upon PDAP1 deficiency was also observed at the protein level (**Figure 4.27D**).

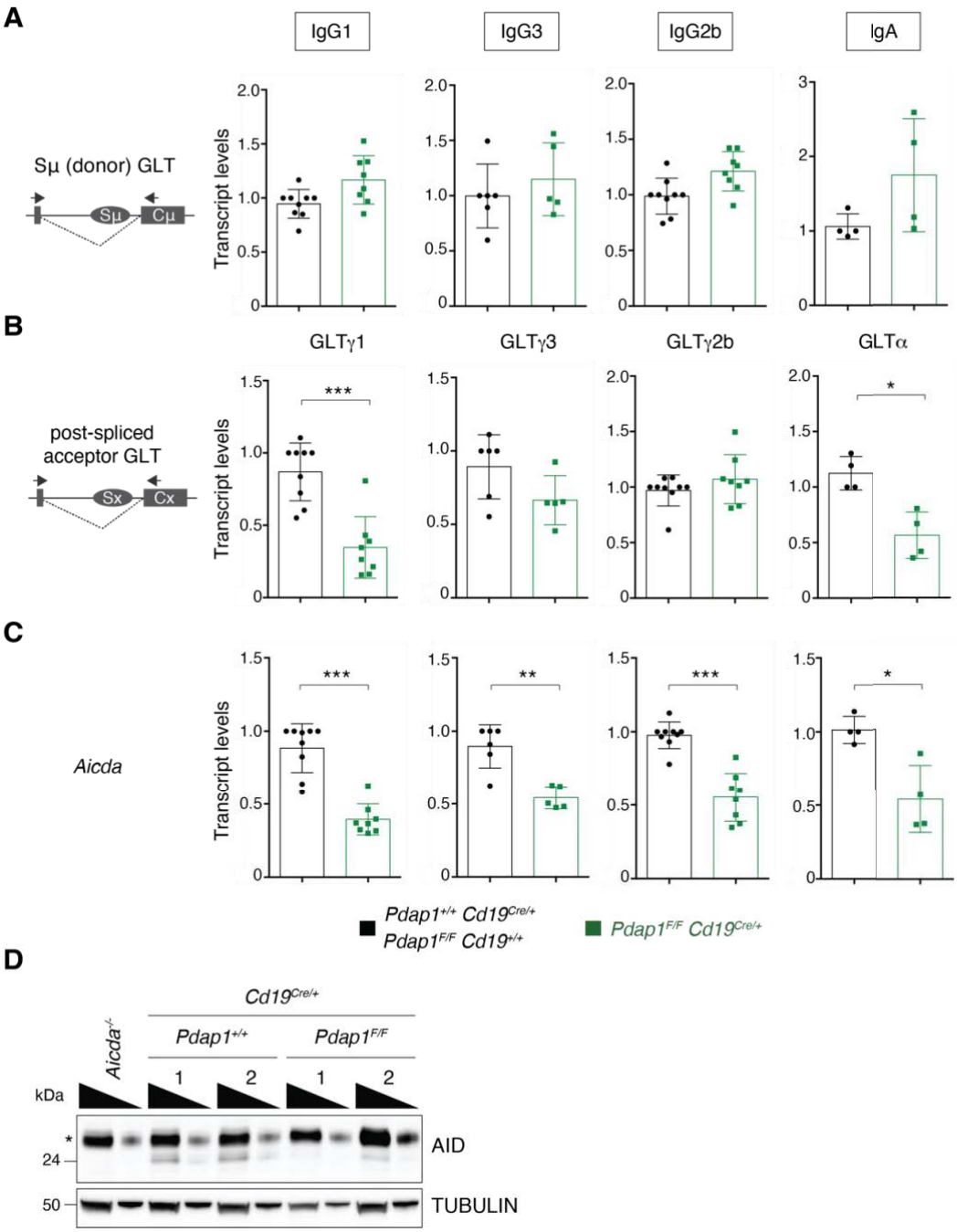

Figure 4.27. PDAP1 is required for efficient AID expression.
A-C. Graphs summarizing RT-qPCR analysis for GLT levels of post-spliced Igμ (**A**) and Igγ1, Igγ3, Igγ2b and Igα (**B**) and *Aicda* mRNA levels (**A**) in primary B-lymphocytes stimulated to

undergo CSR to IgG1 (LPS and IL-4), IgG3 (LPS), IgG2b (LPS, TGFβ and BAFF) and IgA (LPS, TGFβ, BAFF and IL-5) for 48 h. Significance was calculated with the Mann-Whitney U test and error bars represent SD. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001. **D.** WB analysis of WCE of splenocytes of the indicated genotypes cultured in the presence of LPS and IL-4 for 48 h. Triangles represent threefold dilution.

In order to confirm that the CSR defect of *Pdap1*$^{F/F}$*Cd19*$^{Cre/+}$ B cells is due to reduced AID expression, AID was overexpressed in these cells to assess if CSR levels could be rescued. Specifically, primary B cell cultures were infected with a plasmid containing the cDNA of *Aicda* fused to the hormone-binding domain of the modified estrogen receptor (AID-ER). Upon tamoxifen treatment (4-hydroxytamoxifen, 4-HT), AID translocates into the nucleus to mediate CSR[84]. Considering that AID targeting is dependent on GLT, this experiment was performed both in IgG1 (where a decrease in Igγ1 transcription was observed) and IgG2b (where GLT was not affected) conditions (**Figure 4.27B**). Moreover, AID-deficient primary B cells were included as a positive control, which underwent CSR only when infected with the AID-ER plasmid and treated with tamoxifen (**Figure 4.28**). The CSR defect in *Pdap1*$^{F/F}$*-Cd19*$^{Cre/+}$ cells was rescued to similar levels as in AID-ER-reconstituted *Aicda*$^{-/-}$ splenocytes when inducing AID overexpression in the nucleus of IgG2b- but not of IgG1-activated cells. This result could be explained by the differential effect of PDAP1-deficiency in Igγ1 and Igγ2b GLT (**Figure 4.27B**). Altogether, these data indicated that PDAP1 deficiency impairs AID expression and, consequently, antibody isotype differentiation.

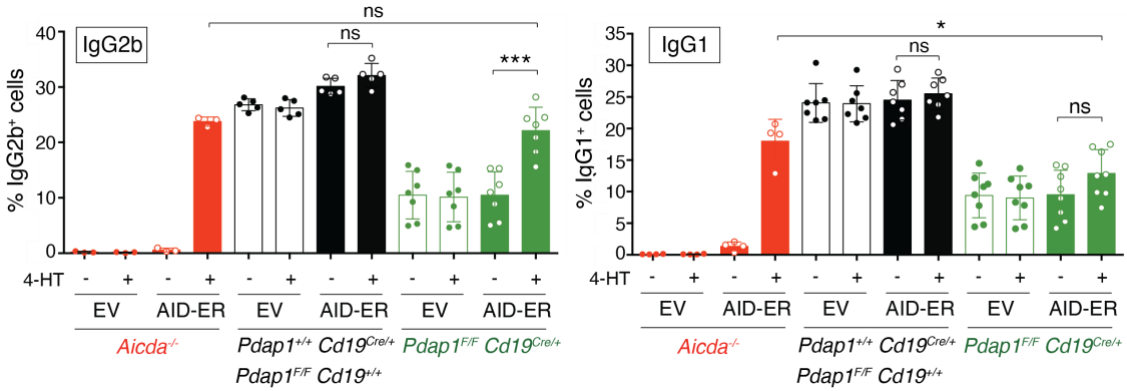

Figure 4.28. AID overexpression rescues IgG2b CSR.
A-C. Graph summarizing the percentage of IgG2b (left) and IgG1 (right) positive cells after retroviral infection of splenocytes of the indicated genotypes with empty vector (EV) or AID-ER (AID cDNA fused to the estrogen receptor) constructs. Cells were stimulated to undergo CSR to for 96 h. Tamoxifen (4-HT) treatment was done for 24 h. Data are representative of four independent experiments with at least four mice per genotype. Significance was calculated with the Mann-Whitney U test and error bars represent SD. *p ≤ 0.05, ***p ≤ 0.001.

4.3.4. PDAP1 supports efficient SHM of Ig loci

SHM is a physiological process that is also dependent on AID and *Aicda*$^{+/-}$ mice exhibit defects in both, CSR and SHM[255,256]. In order to assess if the reduction in AID expression observed in PDAP1-deficient cells also leads to a decrease in SHM, the number of mutations in the *Igh* and *Igk* loci were quantified. Specifically, GC B cells from Peyer's Patches of *Pdap1*$^{F/F}$*Cd19*$^{Cre/+}$ and control aged mice were isolated and the intronic regions downstream J_H4 and $J_\kappa5$ were amplified by PCR and sequenced[259,260]. *Pdap1*$^{F/F}$-*Cd19*$^{Cre/+}$ cells showed a significant reduction in the number of mutated sequences and mutation frequency as compared to controls, both in the J_H4 and $J_\kappa5$ downstream intronic regions. PDAP1-deficient GC B cells presented an increase in low-mutated sequences and a decrease in high mutation rates (more than 20) (**Figure 4.29A-D**). Furthermore, PDAP1 did not affect the distribution of nucleotide substitutions that result from AID function, confirming the role of PDAP1 in supporting AID expression rather than processing the processing of AID-induced lesions (**Fig. 21E-F**). These data evidenced that PDAP1 supports physiological levels of SHM of heavy and light chain *Ig* loci. Conclusively, PDAP1 is required for efficient antibody diversification by SHM and CSR in mature B-lymphocytes.

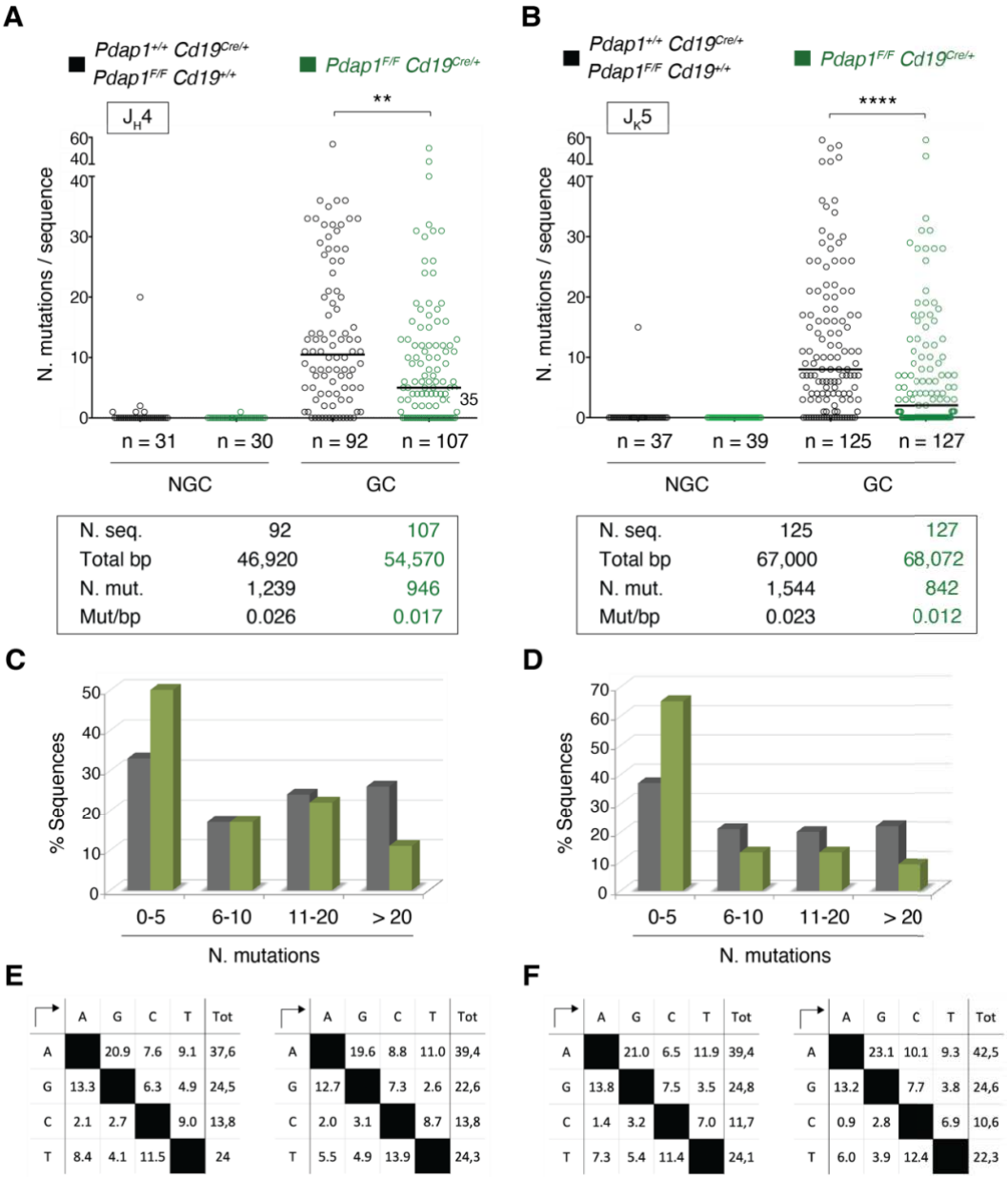

Figure 4.29. AID is required for efficient SHM at Ig loci.

A-B. Up: Graphs summarizing the number of mutations in the 3'J_H4 intronic region over 510bp sequence (**A**) and in the 3'J_K5 intronic region over 536bp sequence (**B**). Genomic DNA from non-germinal center (NGC) and germinal center (GC) B cells from the Peyer's Patches of un-immunized aged mice was amplified, cloned and analyzed (four mice per genotype). Significance was calculated with the Mann-Whitney U test. **p ≤ 0.01, ****p ≤ 0.0001. Bottom: Tables listing the total number and length of analyzed sequences, as well as the number of mutations and mutation frequency at 3'J_H4 (left) and 3'J_K5 (right) intronic regions. (**C-D**) Graphs summarizing the percentage of sequences from panels (A) and (B) presenting the indicated number of mutations in 3'J_H4 (**C**) and 3'J_K5 (**D**). (**E-F**) Profiles of nucleotide substitutions at 3'J_H4 (**E**) and 3'J_K5 (**F**) introns.

4.3.5. PDAP1 is dispensable for the post-transcriptional regulation of *Aicda* mRNA

AID expression is tightly regulated at the level of transcription, post-transcription, post-translational modifications, cellular localization, and targeting to the S regions in order to prevent oncogenic chromosomal translocations (*see section 1.8.2*)[20,168]. Indeed, only a minimum amount of AID protein enters the nucleus upon activating conditions[178–180]. PDAP1 locates in the cytoplasm in primary B cells (data not shown) and, therefore, it might exert a post-transcriptional effect on *Aicda* mRNA. Moreover, since PDAP1 has been recently identified as an RNA binding protein, it could be associating to *Aicda* transcript, modulating AID expression[242–246]. To test this hypothesis, PDAP1-deficient and control splenocytes were isolated and cultured in the presence of Actinomycin D to inhibit *de novo* transcription. Alterations in mRNA capping, adenylation and stability would result in an accelerated mRNA decay, which can be quantified by qRT-PCR[264]. Cells were activated to undergo CSR to IgG1 and IgG2b for 48 h and then collected after 1, 2, 3 and 4 h post-Actinomycin D treatment. Although PDAP1-deficient cells showed a ~50% reduction in *Aicda* mRNA consistent with previous results (**Figure 4.27**), its decay rate was comparable to control cells (**Figure 4.30A-B**). This result suggested that PDAP1 is dispensable for *Aicda* mRNA stability, capping and adenylation.

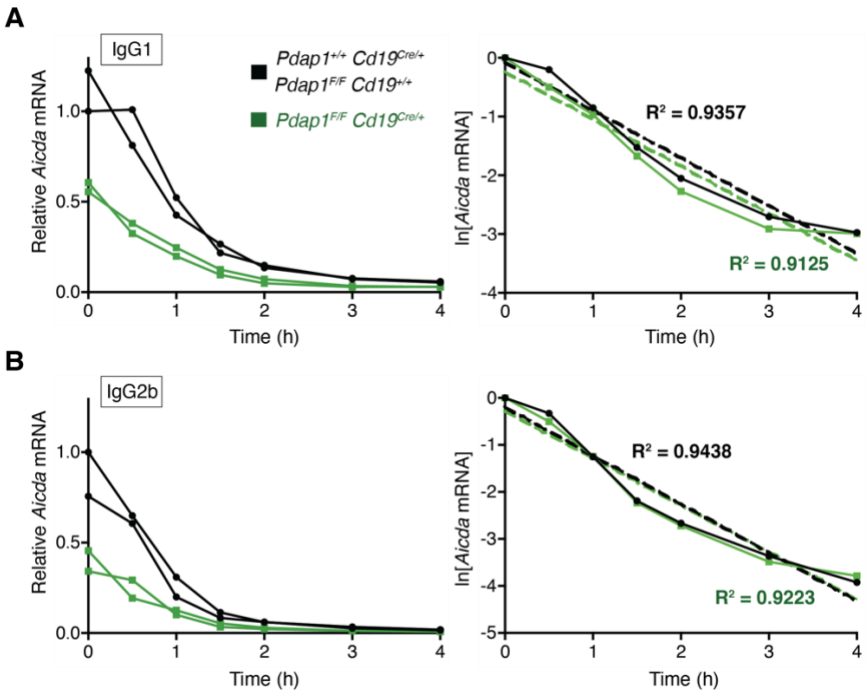

Figure 4.30. PDAP1 is dispensable for *Aicda* mRNA stability.
A-B. Left: Graphs summarizing RT-qPCR analysis for *Aicda* mRNA levels in splenocytes cultured in the presence of LPS and IL-4 (**A**) or LPS, TGFβ and BAFF (**B**) for 48 h and treated with Actinomycin-D over the indicated times. Bottom: Linear regression analysis of data shown on the left panel as ln [RNA] versus time of Actinomycin-D treatment.

Importantly, several non-functional AID isoforms have been described in literature, being the full-length transcript the only splice form encoding a functional AID protein[169,174,175]. Consequently, PDAP1 could be regulating AID splicing, affecting its expression. To test this, RNA-sequencing was performed in splenocytes stimulated to undergo CSR to IgG1 and IgG2b for 48 h and the differential exon usage of *Aicda* mRNA in PDAP1-deficient cells was compared to control cells. This method allows for the quantification of exon expression and identification of different isoforms in RNA-seq samples[206]. As an example and positive control of this analysis, the exon usage of *Pdap1* mRNA was also determined. *Pdap1* exon 2 expression was decreased in *Pdap1*$^{F/F}$-*Cd19*$^{Cre/+}$ cells, which agreed with the conditional mouse model strategy (**Figure 4.22** and **Figure 4.31A**). Contrary, no difference in *Aicda* exon usage was observed upon PDAP1 deficiency, independently of the stimulation conditions (**Figure 4.31B**). Moreover, the unspliced *Aicda* mRNA was also quantified by RT-qPCR, comparing PDAP1-deficient and control cells. As for the spliced *Aicda* mRNA, *Pdap1*$^{F/F}$-*Cd19*$^{Cre/+}$ cells exhibited a significant reduction in the primary transcript (**Figure 4.31C**). Therefore, these data proved that PDAP1 is dispensable for the post-transcriptional regulation of *Aicda* mRNA.

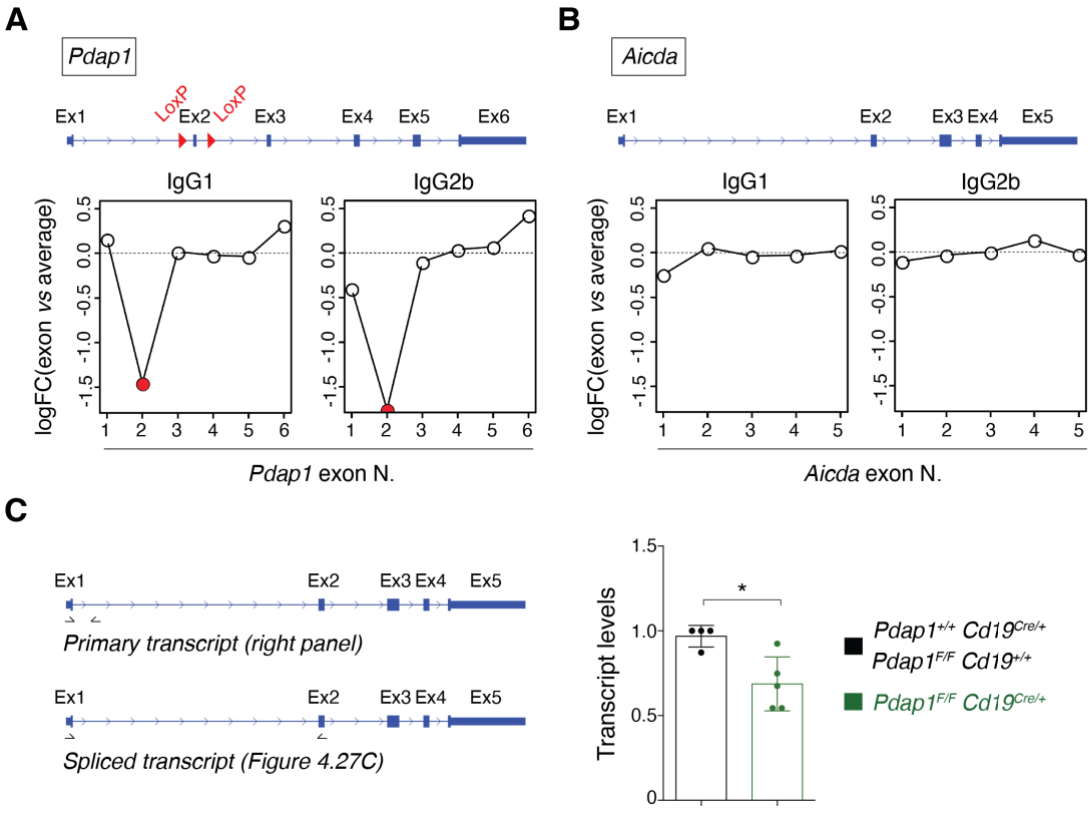

Figure 4.31. PDAP1 is dispensable for *Aicda* mRNA splicing.
A-B. Analysis of exon usage of *Pdap1* (**A**) and *Aicda* (**B**) genes in splenocytes activated to undergo CSR to IgG1 (LPS and IL-4) and to IgG2b (LPS, TGFβ and BAFF) for 48 h. A schematic

representation of each locus is provided above. logFC, log fold change. **C.** RT-qPCR analysis for *Aicda* primary/unspliced transcript in splenocytes of the indicated genotypes stimulated with LPS for 48 h. Data is representative of at least four mice per genotype. One $Cd19^{Cre/+}$ mouse within each experiment was assigned an arbitrary value of 1.

4.3.6. PDAP1 deficiency leads to activation of the integrated stress response in primary B cells

To further characterize the molecular mechanism of PDAP1 in antibody isotype differentiation, the transcriptome of $Pdap1^{F/F}$-$Cd19^{Cre/+}$ splenocytes was compared to $Cd19^{Cre/+}$ control. Since a differential effect in the acceptor GLT was observed in PDAP1-deficient cells depending on the activating conditions, RNA-seq was performed on primary B cell cultures treated with LPS and IL-4 (IgG1) and LPS, BAFF and TGFβ (IgG2b). Importantly, cells were collected 48 h post activation because most of the changes in the transcriptome happen early after CSR induction and before the BCR is expressed in the cell surface[265]. To identify the differentially expressed genes, the significance level of false discovery rate (FDR) was set to < 0.05. The number of differentially regulated genes was higher in LPS-IL-4-stimulated cultures than in LPS-BAFF-TGFβ-activated cells (1227 *versus* 173). Furthermore, the number of genes up-regulated in $Pdap1^{F/F}Cd19^{Cre/+}$ was considerably higher than the down-regulated ones in LPS-BAFF-TGFβ-stimulated cultures (124 up- *versus* 49 down-regulated), but was evenly distributed among the two categories in the LPS-IL-4 stimulation condition (561 up- *versus* 666 down-regulated) (**Figure 4.32**).

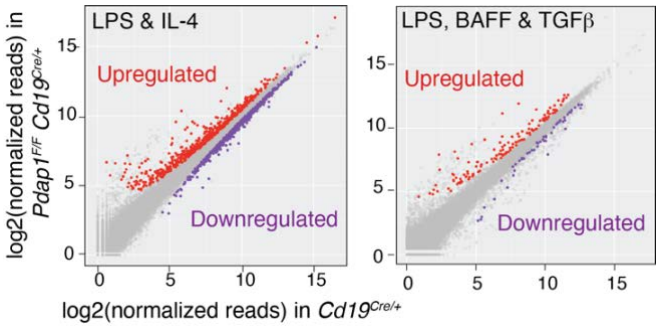

Figure 4.32. Transcriptome analysis of PDAP1-deficient cells.
Scatterplots of gene expression in $Pdap1^{F/F}Cd19^{Cre/+}$ *versus* control ($Cd19^{Cre/+}$) mature B cells stimulated with LPS and IL-4 (left) or LPS, BAFF, and TGFβ (right). Genes with an adjusted p-value (FDR) < 0.05 that are up- or down-regulated in $Pdap1^{F/F}Cd19^{Cre/+}$ cells are highlighted in red or purple, respectively. Data summarizes results from three mice per genotype per stimulation condition and is presented as log2(RPM) (reads per gene per million mapped sequence reads) values.

PDAP1-deficient cells showed a reduction in *Aicda* and *Ighg3* genes in both stimulation conditions, which is in agreement with previous RT-qPCR results (**Figure 4.27B** and

Figure 4.33B). Moreover, pathway enrichment analysis of the upregulated genes in *Pdap1*$^{F/F}$*Cd19*$^{Cre/+}$ B cells in both conditions were related to metabolic pathways of amino acid and aminoacyl-tRNA biosynthesis. Indeed *Asns*, *Cth*, and *Psat1* and *Psph* (required for the synthesis of asparagine, cysteine and serine, respectively) were found among the top upregulated genes together with several amino acid transporters (*Slc1a4*, *Slc7a3* and *Slc6a9*) upon PDAP1-deficiency (**Figure 4.33A**). These "metabolic signature" is characteristic of cells that undergo amino acid deprivation and activate the integrated stress response (ISR). Accordingly, ATF4, which is a main effector of the ISR, was also significantly upregulated in *Pdap1*$^{F/F}$*Cd19*$^{Cre/+}$ B cells, both at the transcript and protein level (**Figure 4.33B-C**). Importantly, the ATF4 pathway has been shown to be suppressed in B-lymphocytes and induced only after activation. This was also observed in *Cd19*$^{Cre/+}$ cells after 72 and 96 h of cytokine stimulation. In contrast, *Pdap1*$^{F/F}$*Cd19*$^{Cre/+}$ splenocytes exhibited an increased ATF4 expression even in resting (unstimulating) conditions. Therefore, all these data indicated that PDAP1 deficiency causes activation of the ISR in B cells. What are the implications of this activated ISR and whether its induction is a direct consequence of PDAP1 depletion or if this factor is exerting an independent function that promotes cellular stress needs to be determined with further experiments.

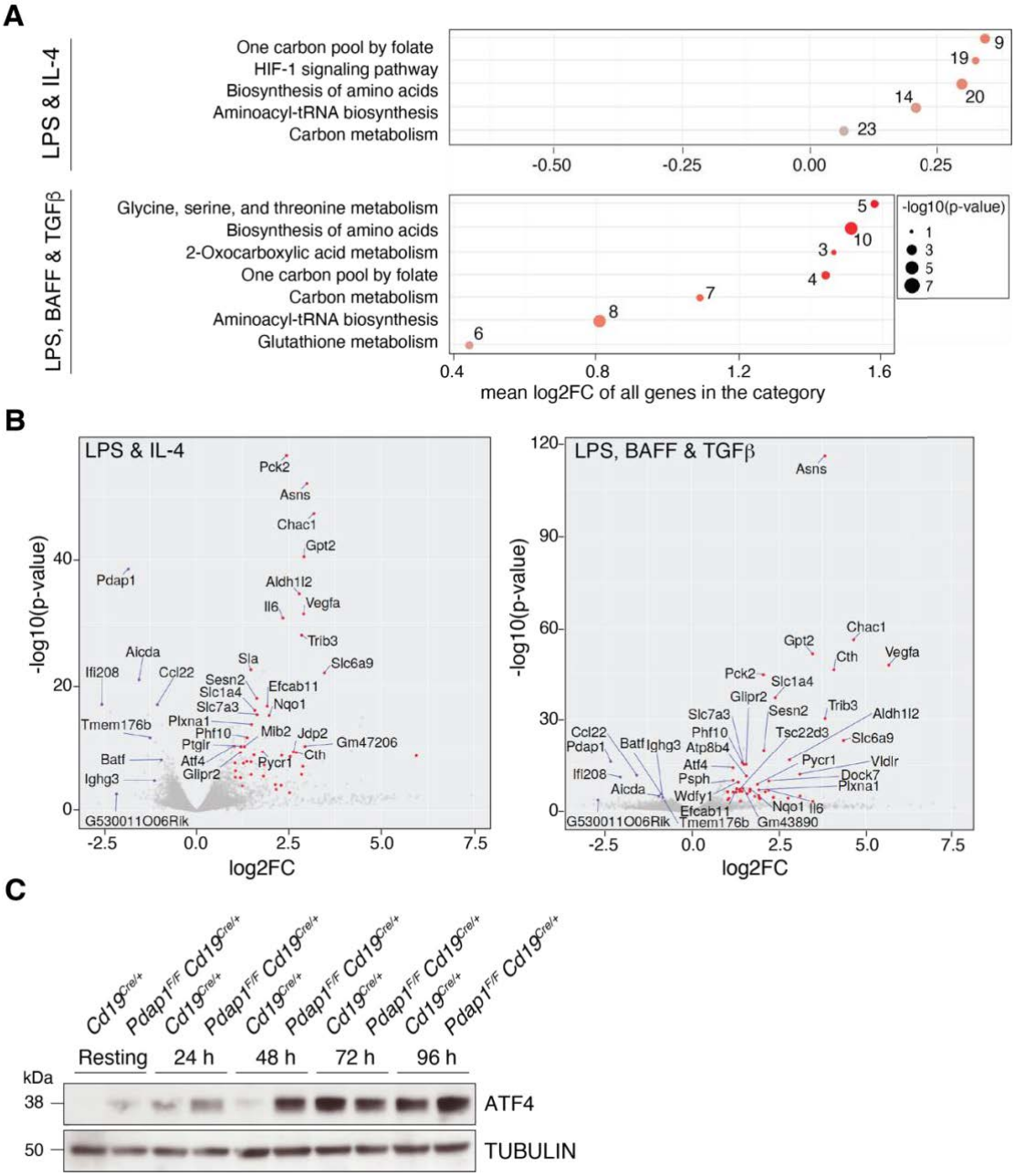

Figure 4.33. PDAP1 deficiency activates the integrated stress response.
A. Pathway enrichment analysis (KEGG pathways) of the differentially regulated genes (FDR < 0.05) from Figure 4.32. The number of differentially regulated genes in each category is indicated. **B.** Volcano plots displaying differentially expressed genes between control and Pdap1-deficient splenocytes activated with LPS and IL-4 or (left) or LPS, BAFF, and TGFβ (right). The red and purple dots represent transcripts up- and down-regulated in *Pdap1*$^{F/F}$*Cd19*$^{Cre/+}$ cells, respectively, with FDR < 0.05 and expression fold change > 2 (up-regulated) or 1.7 (down-regulated) in both stimulation conditions. The names of the down-regulated and the 30 most significantly up-regulated genes within each stimulation condition are indicated in each graph. The fold change threshold for down-regulated genes was set to 1.7 to include genes yielding a biologically relevant effect even with less pronounced variations in expression levels (e.g. *Aicda*, haploinsufficient gene). **C.** Representative WB analysis of splenocytes of the indicated genotypes before (resting B cells) and after 24 to 96 h activation with LPS and IL-4. Data is representative of two mice per genotype.

5. DISCUSSION

Upon antigen encounter in the secondary lymphoid organs, mature naïve B cells undergo SHM and CSR in order to produce antibodies that have high specificity for the antigen and different effector properties. These two processes are therefore essential for the establishment of an effective immune response and alterations in any of the two could cause immunodeficiency and/or cancerous genomic rearrangements. Although the molecular aspects of SHM and CSR have been extensively studied, the full picture of factors regulating antibody diversification in mature B cells is still missing. The loss-of-CSR assay developed in this study represents a useful tool for the identification of novel players that contribute to mediate an effective immune response in B-lymphocytes. By applying this screening method, ZMYND8 and PDAP1 have been identified as novel CSR factors. Their contribution to *Igh* gene diversification has been dissected, thus providing new insights in the mechanisms that contribute to the generation of the antibody diversity.

5.1 Advantages and caveats of the loss-of-CSR assay

The setup of the loss-of-CSR assay allowed for rapid, robust and functional screening of potential candidates involved in antibody isotype differentiation. The characterization of multiple negative and positive controls ensured the establishment of a large dynamic range to perform the screen. Moreover, CH12 cells are an ideal model system for this purpose not just because it is the only available cell line that undergoes CSR with high efficiency, but also because these cells divide every 8-12 h[189,216]. The optimized conditions led to the identification of ZMYND8 and PDAP1 as positive hits in two independent screening rounds. Their role in CSR was later validated in CH12 KO clonal derivatives (ZMYND8-KO and PDAP1-KO) and in conditional mouse models (*Zmynd8*$^{F/F}$-*Cd19*$^{Cre/+}$ and *Pdap1*$^{F/F}$-*Cd19*$^{Cre/+}$), which prove the robustness of the screen. Furthermore, ZMYND8 and PDAP1 were shown in this study to have different functions during CSR, evidencing the potential of the screen to reveal novel factors involved in different steps of antibody isotype differentiation. Of note, the majority of the tested candidates did not show any reduction in CSR and this could be due to several reasons. Firstly, the gRNAs used were designed using robust software tools that predicted off-target effects and generated gene/sequence-specific gRNAs (www.crispr.mit.edu[194] and CrispRGold[195]). However, these gRNAs were not experimentally tested for their targeting efficiency or off-target effects. To increase the chances of causing indels at the desired locus, 3-6 gRNAs per candidate were designed and electroporated together as a pool in

equal ratio. Nevertheless, this does not exclude the option that a negative result in the screen could be due to inefficient gRNA targeting. Secondly, it is possible that CRISPR/Cas9 mediated the deletion or truncation of a factor that is required for B cell proliferation. Consequently, positively targeted cells would have died, and the non-targeted cells would have proliferated and led to physiological levels of CSR. Lastly, each factor displays its own turnover ratio. This implies that, even if the desired locus was efficiently targeted, the previously translated protein might had not been fully degraded when cells were activated to undergo CSR. Therefore, all these considerations imply that the negative hits in the screen could still be factors involved in antibody isotype differentiation.

5.2 ZMYND8 function in *Igh* gene diversification

5.2.1 ZMYND8-RIF1 interaction

RIF1 has been shown to play a role in the repair phase of CSR by preventing DNA end resection and, therefore, supporting NHEJ. Consequently, *Rif1*$^{-/-}$ B-lymphocytes are unable to repair AID-induced breaks at the *Igh* locus. This leads to a major reduction in productive CSR (80-85%) but also to the accumulation of chromosome breaks, translocations and genomic instability[96,97,100,137,138]. Considering the relevance of RIF1 function in CSR, a SILAC-based approach was set up in splenocytes stimulated *ex vivo* for the identification of novel interactors that could be regulating antibody isotype differentiation (**Figure 4.2**). In this context, ZMYND8 was found to be a RIF1 interacting partner and later shown to be required for efficient CSR (**Figure 4.4, 4.7 and 4.8**). Importantly, co-immunoprecipitation studies proved that, unlike phosphorylated 53BP1-RIF1 association, this interaction is DNA-damage-independent (data not shown). ZMYND8 was dispensable for DSB repair, independently of the model system used (CH12, primary B-lymphocytes or iMEFs) and the source of DNA breaks (AID, Cas9, IR or PARPi) (**Figure 4.11-13**). Furthermore, RIF1-deficiency did not lead to an increase of RNAPII loading at the 3' regulatory region as it was observed in *Zmynd8*$^{-/-}$ cells (**Figure S3**). Altogether, these findings indicate that ZMYND8 and RIF1 clearly have independent functions during antibody isotype differentiation. Consequently, it is possible that their interaction could be stochastic or not have any implication during CSR. On the other hand, RIF1 and ZMYND8 could potentially function together in CSR or other physiological processes independently from their role in DSB repair and 3'RR transcriptional regulation, respectively. In this context, it has been recently shown that 53BP1-dependent RIF1 protection of DSB DNA ends is required but not essential for CSR to happen. Specifically, *Trp53bp1*$^{-/-}$ splenocytes complemented with a mutant

version of 53BP1 that is unable to recruit RIF1 and prevent DNA end resection upon DSB formation can undergo CSR to a considerable extent[266]. Also, an independent study showed that inhibition of CtIP, an endonuclease that resects DNA ends, does not rescue the CSR defect observed upon RIF1 deficiency[96]. Lastly, several articles have pointed out the role of RIF1 in transcriptional regulation, both in mammals and yeast[152–155,267–271]. Altogether, these data support the interesting and novel idea that RIF1 presents an additional function in antibody isotype differentiation prior to DSB formation. In this regard, there might be other potential RIF1 interactors identified in the I-DIRT that may contribute to RIF1 additional function(s) in CSR. Moreover, it would be interesting to elucidate if RIF1-deficient splenocytes show a decrease in the number of AID-induced mutations. Another interesting approach to further study RIF1-ZMYND8 interaction would be to reconstitute ZMYND8-deficient primary B cells with the full length ZMYND8 protein together with an array of mutant forms, measuring CSR levels and performing co-immunoprecipitation. This would characterize the domain of ZMYND8 that is required for its interaction with RIF1. Furthermore, this experiment would determine if any of the ZMYND8 mutants is able to rescue CSR. Of note, RIF1 is a large protein (2426 amino acids) compared to ZMYND8 (1235 amino acids), impeding the packaging of RIF1 coding sequence into a retroviral vector and its transduction efficiency in B cells.

5.2.2 ZMYND8 role in DNA repair

Human ZMYND8 has been previously reported to support the recruitment of the NuRD complex to sites of DNA damage within transcriptionally active chromatin to promote transcriptional repression and subsequent repair by HR[229–233]. CSR is mediated by c-NHEJ, so it is not surprising that ZMYND8 did not play a role in DSB repair in this context. However, no significant defect in cell survival or increased chromosomal aberrations were observed when comparing *Zmynd8*$^{-/-}$ to *wt* cells upon PARPi treatment (**Figure 4.12B-C**). Importantly, the reduction in HR upon ZMYND8 deficiency previously described was partial and only observed upon simultaneous laser micro-irradiation (which induces massive DNA damage in a specific location) and PARP inhibition[229,232]. Indeed, the NuRD complex can still associate to target genes in the absence of ZMYND8. Recruitment of the different NuRD complex subunits to sites of DNA damage and subsequently DSB repair by HR are only partially reduced in ZMYND8-deficient cells[232]. Therefore, ZMYND8 seems to have a minor role in DNA repair that may not be detected by the experimental conditions used in this study.

5.2.3 Transcriptional regulation of the *Igh* 3' RR

This study evidences that increased transcription of the 3'RR has a negative effect in CSR. The murine 3'RR is a palindromic super-enhancer of 28 kb composed of four core enhancers. The 3'RR is involved in late B cell development and during the early steps of antibody isotype differentiation, specifically in GLT, chromatin accessibility, RNAPII loading and pausing and AID recruitment[134]. Moreover, this super-enhancer also promotes SHM at the variable region of the *Igh* locus (*see section 1.6*)[24]. Interestingly, ZMYND8 binds to the 3'RR and regulates its transcriptional status, acting as a repressor. ZMYND8-deficient mice showed increased RNAPII loading and enhancer transcription of the 3'RR (**Figure 4.17-18**), decreased CSR and GLT of acceptor S regions (**Figure 4.7, 4.8 and 4.16**), and reduced SHM exclusively in the heavy chain locus (**Figure 4.19**). Therefore, some key aspects of previously published work about the 3'RR functions correlate with the described observations upon ZMYND8 deficiency and will be discussed more in more detail below.

<u>B cell development</u>

The 3'RR and specifically the hs4 core enhancer have been shown to regulate late B cell development, promoting the transition from immature to mature B cells and, supporting efficient Ig expression and B cell differentiation into the marginal zone compartment[22–25,129,132]. ZMYND8 associates with the 3'RR and significantly represses RNAPII loading and transcription of hs1,2 and hs3b core enhancers (**Figure 4.17-18 and S2**). *Zmynd8*$^{F/F}$*-Cd19*$^{Cre/+}$ mice did not show any major defect in B cell development, which is in agreement with the fact that the hs4 region (crucial for B cell development) is unaffected in these mice (**Figure 4.6**). However, it is important to consider that Cre-mediated *Zmynd8* deletion is not fully efficient and this may impede to observe a phenotype in B cell development. Since *Zmynd8*$^{-/-}$ mice are likely lethal, this hypothesis could be tested by analyzing B cell compartments/development in *Zmynd8*$^{F/-}$*-Cd19*$^{Cre/+}$ mice.

<u>Somatic Hypermutation</u>

The palindromic module of the 3'RR (hs3a-hs1,2-hs3b) regulates SHM of the heavy but not of the light chain[129]. Consequently, SHM of the *Igh* locus is severly defective in mice bearing the following mutations: Δ3'RR (complete absence of the super-enhancer), Δhs3a-hs1,2, Δhs3a-hs1,2-hs3b, ΔIRIS (altered palindromic structure of hs3a and hs1,2) and core-3'RR (lack of the packaging DNA that links the 4 core enhancers) (**Figure 1.7**) (*see section 1.6*)[24,128,129,135]. Likewise, ZMYND8 promotes SHM exclusively in the heavy

chain locus, supporting the idea that this factor plays a central role in the regulation of the 3'RR function in mature B-lymphocytes (**Figure 4.19**).

Germline transcription of S regions

While deletion of individual core enhancers does not affect CSR, removal of the entire 3'RR completely abrogates GLT and antibody isotype differentiation to all isotypes[22–24,26,132–134]. Δ3'RR partially reduces transcription of the donor Sμ region and totally abolishes the expression of acceptor S regions. Thus, it has been hypothesized that this super-enhancer regulates GLT and accessibility of the acceptor S regions, being Sμ less 3'RR dependent[134]. ZMYND8 deficiency leads to decreased GLT of acceptor S regions, while transcription of the donor Sμ region is unaffected (**Figure 4.15**). This phenotype was also observed in cells presenting a combined deletion of hs3b and hs4 core enhancers (Δhs3b-hs4) or a replacement of hs1,2 with an actively transcribed neomycin cassette. In both cases, splenocytes showed diminished GLT of several acceptor S regions while preserving WT levels of Sμ transcription. However, these mutations did not affect IgG1 surface expression or γ1 GLT, which were significantly reduced in ZMYND8-deficient cells (**Figure 4.7A and 4.16**)[25,26].

Importantly, the reduced GLT observed in 3'RR-deleted splenocytes correlated with a decrease in both, paused and elongating phosphorylated RNAPII (RNAPII-S5 and RNAPII-S2, respectively) at acceptor S regions[134]. This phenotype was different in ZMYND8-deficient cells, which showed reduced levels of γ1 GLT but not of RNAPII-S5 at Igγ1 region. Of note, the decrease in γ1 GLT was determined by RT-qPCR (**Figure 4.16**) and by GRO-seq analysis (global run-on sequencing), which measures nascent RNA species in isolated nuclei (data not shown). These two assays also confirmed the increased enhancer transcription of the 3'RR in ZMYND8-deficient splenocytes, which correlated with an increase in RNAPII-S5, compared to control cells. Therefore, transcription of acceptor S regions is reduced upon *Zmynd8* deletion but RNAPII loading in these regions is unaffected. One possible explanation is that the increase in RNAPII loading at the 3'RR in ZMYND8-deficient cells leads to local competition of *Igh* transcription factors, biasing the positioning of the transcriptional machinery towards the 3'RR and reducing its activity at the acceptor S regions. Alternatively, the rate of elongating RNAPII (RNAPII-S2) at acceptor S regions upon ZMYND8 deficiency could be reduced, explaining the decrease in GLT while preserving the amount of loaded polymerase (RNAPII-S5). In agreement with this idea, ZMYND8 has been shown to interact and recruit P-TEFb (positive transcription elongation factor b) that facilitates the release of paused RNAPII by mediating its phosphorylation at Ser2. Accordingly, ZMYND8-deficient cells present a decrease in p-TEFb recruitment to target genes[226].

Therefore, it would be necessary to perform a RNAPII-S2 ChIP-seq to further dissect the molecular mechanism by which ZMYND8 mediates its function in GLT regulation and CSR.

Igh chromatin architecture

Chromatin organization of the heavy chain is crucial for efficient GLT and CSR to happen. Mature resting B cells exhibit a looped *Igh* locus, characterized by the contact of Eμ and the 3'RR. Cytokine stimulation leads to inducible transcription of a specific GLT promoter that is recruited to the Eμ-3'RR loop base, allowing transcription of the S region, AID targeting and CSR[114,115]. This architectural configuration is dependent on the 3'RR, as combined deletion of core enhancers hs3b and hs4 impairs the loop formation in between Eμ and the 3'RR[115]. ZMYND8-deficiency leads to transcriptional de-regulation of the 3'RR, which could lead to inefficient *Igh* looping. This would explain why increased transcription of this super-enhancer has a negative effect on GLT of acceptor S regions. Circularized chromosome conformation capture combined with sequencing (4C-seq) of ZMYND8-deficient splenocytes in resting conditions and also stimulated with LPS and IL-4 for 48 h did not show any difference in the interactions within the *Igh* locus compared to WT cells (data not shown). This result indicates that ZMYND8 is dispensable for the chromatin architecture of the *Igh* locus. Therefore, the transcriptional de-regulation of the 3'RR and GLT reduction caused by ZMYND8 deficiency does not correlate with a defect in *Igh* looping. Of note, the 4C-seq experiment was performed by using a region within the 3'RR as an anchor, and it has been previously discussed that the resolution of this assay can vary depending on the bait used[272]. Consequently, it would be interesting to repeat this analysis, using different bait locations in parallel and compare the results. An alternative could be anchoring Eμ, which has been widely used in previous studies to analyze the chromatin configuration of the *Igh* locus[115,118,273]. Additionally, it is possible that ZMYND8-deficiency leads to subtle chromatin architectural changes that may not be detectable by 4C but still able to disrupt the 3'RR function.

5.2.4 ZMYND8 as a transcriptional repressor and activator

Apart from being a transcriptional repressor, human ZMYND8 has been also shown to bind and recruit P-TEFb to transcription start sties (TSS) of active genes[226]. P-TEFb function is crucial to promote the transition from paused to elongating state of RNAPII and, therefore, reduced P-TEFb recruitment correlates with decreased transcription of target genes[274]. Moreover, human ZMYND8 can be found as a monomer or a dimer and it has been suggested to have different roles in transcription depending on its

configuration[226]. Specifically, monomeric ZMYND8 has been suggested to bind CHD4 to allow the NuRD complex recruitment to enhancers and promoters and transiently repress transcription during DNA repair of active genes. On the other hand, human ZMYND8 can also form homodimers, which preferentially associates to P-TEFb and positively regulates transcription[226]. Importantly, human ZMYND8 has been also reported to repress transcription at enhancers and super-enhancers through its interaction with KDM5C and independently of the NuRD complex in human cells[227]. Consequently, it is likely that ZMYND8 role in transcriptional regulation could be different depending on the chromatin element it binds and the different factors that is associating to. This implies that ZMYND8 could mediate transcriptional activation of the acceptor S regions while independently repressing the transcription of the 3'RR region. Therefore, it would be interesting to study if murine ZMYND8 binds to DNA elements (promoters and enhancers) within the *Igh* locus in distinctive mono-/oligomeric states and if it is interacting with different partners to promote transcriptional repression or activation. Moreover, in this study ZMYND8 was found to bind the putative enhancer element in between Cγ1 and Sγ2b, Eγ. Although the function of this regulatory element is still elusive, it has been recently shown to be required for efficient CSR to some isotypes[31]. It would be then interesting to further characterize the role of this putative enhancer in mature B cells and the consequences of ZMYND8 binding to this region.

5.2.5 Molecular mechanism of ZMYND8 function in CSR

Taking into account all the phenotypical variances shown by the different mutants of the 3'RR described in the literature and the fact that ZMYND8 preferentially represses the transcription of hs1,2 and hs3b core enhancers, it is not surprising that ZMYND8-deficiency does not completely phenocopy all the aspects characteristic of 3'RR-deleted cells. In this study, a novel function for the chromatin reader ZMYND8 has been described as the first modulator of the 3'RR transcriptional activity. Specifically, ZMYND8 supports GLT of acceptor S regions and CSR *in vivo* as well as efficient SHM of the *Igh* locus, which, as CSR, is also under the control of the 3'RR. ZMYND8 deficiency leads to increased transcription of this super-enhancer and the data favors the hypothesis that this de-regulation is responsible for the decreased transcription of acceptor S regions, CSR and SHM and the of the *Igh* locus (**Figure 5.1**) In agreement with this idea, replacement of single hs1,2 or hs3a or paired hs3b-hs4 core enhancers with a neomycin cassette that is actively transcribed affected GLT and CSR. On the other hand, individual deletion of the same modules did not have any effect on CSR and hs3b-hs4 abrogation resulted in a less severe defect in GLT and CSR than its replacement with a neomycin cassette[25–27]. However, further analyses need to be performed to support this link. A

possibility would be to use a catalytically inactive Cas9 (dCas9) fused to transcriptional effectors that can be guided to the chromatin for silencing or activating the 3'RR. Specifically, dCas9-KRAB could be targeted to the hs1,2 and hs3b regions to repress their overexpression in order to assess whether this would rescue the GLT and CSR defect of ZMYND8-deficient splenocytes. Alternatively, WT cells could be infected with dCas9-VP64 to overexpress the 3'RR and test if this causes a decrease in GLT and CSR as in *Zmynd8* deleted B lymphocytes.

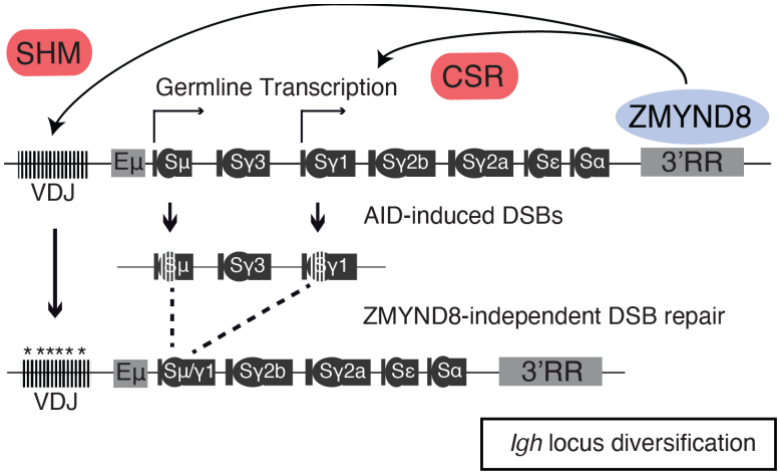

Figure 5.1. Model for ZMYND8 function at the Igh locus.
ZMYND8 binds to and regulates the activity of the 3'RR, controlling CSR and SHM. GLT and CSR to IgG1 are exemplified. Asterisks indicate mutations at the VDJ region as a consequence of SHM.

5.3 PDAP1 role in B cell physiology

5.3.1 PDAP1 regulation of *Aicda* expression and GLT

In this study, PDAP1 was found to support efficient *Aicda* expression since PDAP1-deficient splenocytes showed a ~50% reduction in *Aicda* mRNA levels (**Figure 4.27C-D**). Consequently, CSR and SHM were significantly decreased upon *Pdap1* deletion. These results are in agreement with previously published data revealing that *Aicda* is haploinsufficient. Accordingly, mice harboring only one functional *Aicda* allele show a defect in CSR and SHM[255,256]. Moreover, exogenous induction of AID expression in the nucleus of *Pdap1*$^{-/-}$ cells restored the decreased IgG2b surface expression to physiological levels (**Figure 4.28**). PDAP1 has been previously reported to bind RNA in human cell lines[242–246] and, therefore, it may associate to *Aicda* transcript and regulate its splicing or stability. However, data resulted from this study conclusively indicated that PDAP1 is dispensable for the post-transcriptional regulation of *Aicda* mRNA (**Figure 4.30**). Considering that PDAP1 localizes in the cytoplasm (Human Protein Atlas and data

not shown), it is likely that this factor is indirectly modulating *Aicda* transcription, rather than acting as a transcription factor. In support of this idea, RNA-seq analysis of primary B cell cultures undergoing CSR showed a significant increase of *Trib3* transcript levels upon PDAP1 deficiency (**Figure 4.33B**). TRIB3 has been shown to act as an inhibitor of NF-κB dependent transcription, which is required for efficient AID expression (*see section 1.8.2*)[275–278]. Furthermore, PDAP1-deficient splenocytes displayed reduced *Batf* expression, and *Batf* has been reported to support *Aicda* and GLT of acceptor S regions (**Figure 4.33B**)[279]. This result is in agreement with the fact that *Pdap1* deletion also leads to a decrease in GLT of several acceptor regions compared to control, although transcription of Igγ2b was unaffected (**Figure 4.27B**). These differential effect of PDAP1 in GLT could explain why AID overexpression in PDAP1-deficient cells rescued the CSR defect to IgG2b but not to IgG1 (**Figure 4.28**). Importantly, activation of germline promoters with different cytokine cocktail mixes leads to the recruitment of independent and distinct transcription factors to the *Igh* locus[280,281]. Consistently with this idea, IL-4 activates STAT6 that associates with the Igγ1 and Igγ3 promoters whereas TGFβ stimulates SMADs and RUNX proteins, which promote Igγ2b transcription[63,65]. This implies that PDAP1 could be directly or indirectly regulating GLT of Igγ1 but not of Igγ2b. Although the currently available data suggests that PDAP1 is not a transcription factor, it would be interesting to determine its chromatin association profile in activated primary B cells to confirm this hypothesis. Moreover, it would be helpful to perform PAR-CLIP (Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation) in switching splenocytes in order to identify PDAP1 RNA binding sites and RNA interactome. This analysis would determine if PDAP1 is directly associating to *Aicda* mRNA and/or germline transcripts and would provide new insights to further characterize the role of PDAP1 in antibody isotype differentiation.

5.3.2 PDAP1 and the integrated stress response

PDAP1-deficiency is accompanied by the activation of the integrated stress response (ISR), as determined by RNA-seq signature of activated mature B-lymphocytes (**Figure 4.33**). Importantly, ATF4 expression is induced as a consequence of amino-acid deprivation, oxidative stress and the unfolded protein response (UPR). ATF4 plays a major role supporting the transcription of stress-related genes and factors required for amino acid synthesis and transport, regulating the cell metabolism[282–285]. The UPR is crucial during an immune response for the differentiation of B-lymphocytes into antibody-secreting plasma cells. Specifically, the immunoglobulin heavy and light chains are translocated into the endoplasmic reticulum (ER), where they are assembled to form functional molecules. As antibody production increases, the ER activates the UPR to

augment its protein folding capacity[286–289]. The UPR comprises three branches that involve the activity of independent transducing factors during ER stress, promoting the transcription of UPR target genes. Upon stress, ATF6 translocates into the Golgi apparatus, where it is cleaved and, thus, is able to move to the nucleus and activate genes involved in protein folding. Independently, PERK dimerizes and is phosphorylated, allowing the subsequent phosphorylation of eIF2α and global translation inhibition. As a consequence of PERK and eIF2α phosphorylation, ATF4 translation is increased, which promotes the transcription of CHOP that is involved in apoptosis. Lastly, IRE1 undergoes oligomerization upon stress and cleaves a 26-nucleotide segment from the mRNA of *Xbp1*, a UPR-specific transcription factor. As a result, there is a shift in the reading frame of *Xbp1* and this spliced version of the protein is translocated into the nucleus, where it induces the expression of target genes[283]. Importantly XBP1 is crucial for the expansion of the secretory pathway in plasma cells and *Xbp1* deleted B-lymphocytes show reduced levels of antibody secretion compared to control cells. Moreover, BLIMP-1, encoded by *Prdm1*, has been reported to act upstream of XBP1 during plasma cell differentiation[290,291]. Interestingly, both *Atf4* and *Prdm1* transcripts were upregulated upon PDAP1 deficiency, indicating the potential role of PDAP1 in activating the stress response and plasma cell differentiation. PDAP1-deficient cells might over-activate the UPR, leading to an increase in the number of plasma cells and the levels of secreted antibodies. In agreement with this idea, there was a decrease in the percentage of germinal center (GC) B cells compared to control upon *Pdap1* deletion (**Figure 4.25**). Therefore, it would be interesting to determine the levels of plasma cells in the bone marrow and spleen of PDAP1-deficient splenocytes as well as measuring the percentage of plasmablasts in activated primary B cell cultures. In this context, it would be also necessary to quantify the number of antibody secreting cells by ELISpot (enzyme-linked immunospot) and antibody titers in serum by ELISA. Furthermore, it would be interesting to assess the levels of spliced *Xbp1* transcript and protein in PDAP1-deficient and control cells undergoing CSR. These analyses would elucidate if PDAP1 is playing a role during plasma cell differentiation.

Recent literature show that B cells suffer broad metabolic changes during activation, including the upregulation of oxidative phosphorylation (OXPHOS), tricaboxylic acid cycle (TCA) and nucleotide biosynthesis[292–297]. Stimulated B cells also modify their mitochondrial mass, membrane potential and architecture, changing from few elongated mitochondria with multiple nucleoids to multiple rounded mitochondria with a single nucleoid[296,298]. Moreover, mitochondrial stress has been shown to inhibit components necessary for OXPHOS and to promote the induction of ATF4 together with its

downstream targets, which mediate the ISR. Specifically, ATF4 supports the expression of several amino-acid transporters (SLC6A9, SLC1A4, SLC7A3, SLC1A5, SLC7A11), enzymes involved in the synthesis of amino-acids (ASNS, PSPH, PHGDH, GPT2, PCK2) and enzymes required for mitochondrial one-carbon metabolism and glutathione degradation (ALDH1L2 and CHAC1). This leads to metabolic reprogramming of the cell that ensures adaptation to stress[299]. In agreement to this, ATF4 has been reported to play a major role in metabolic remodeling of activated $CD4^+$ T cells[300]. Importantly, ATF4 also promotes cell death during persistent ISR by activating downstream pro-apoptotic targets, such as CHOP and TRIB3[301–304]. Considering that ATF4 and the majority of these factors are upregulated in stimulated PDAP1-deficient splenocytes, it is possible that PDAP1 supports mitochondrial function, ensuring survival and activity of mature B cells. This hypothesis could explain the decrease in the number of mature splenocytes and the reduction in GC B cells upon *Pdap1* deletion (**Figure 4.23C and 4.25**). However, further studies need to be performed to confirm this hypothesis. Specifically, it would be necessary to determine the rate of mitochondrial respiration (oxygen consumption rate, OCR) and the level of glycolysis (extracellular acidification rate, ECAR) in naïve and activated PDAP1-deficient cells. These parameters, together with an additional array of measurements that can be obtained by using the Seahorse Analyzer could provide an insight into the mitochondrial functioning of *Pdap1* deleted B cells, compared to control. Furthermore, it would be worth quantifying the level of apoptosis and cell number of PDAP1-deficient and control splenocytes following activation to undergo CSR as well as characterizing the expression and signaling of the effectors of the ISR.

Conclusively, activation of the ISR in PDAP1-deficient cells could have different outcomes in B cell fate and function. Further analyses need to be performed to determine what are the cause/s and consequences of the induction of this stress response and what is the link to the reduced number of GC B cells, *Aicda* mRNA expression, CSR and SHM observed upon PDAP1-deficiency.

CONCLUSION AND FUTURE OUTLOOK

Altogether, the identification and characterization of novel CSR factors provide valuable information about the requirements for the establishment of an efficient immune response and further increase our knowledge about B cell development and functioning. Apart from the points discussed above, this study opened several questions and new lines of research:

- Investigating the potential novel role of RIF1 in CSR before DSB repair.
- Further defining the mechanism of ZMYND8 function in the *Igh* locus transcription.
- Studying the implications of the 3'RR overactivation in B cell development and function.
- Understanding the role of PDAP1 in CSR.
- Characterizing the consequences of the ISR induction during B cell activation.

SUPPLEMENTARY FIGURES

Sμ-Sγ1	$Cd19^{Cre/+}$ (30N)	$Zmynd8^{F/F}$ $Cd19^{Cre/+}$ (29N)
Mouse #1	TAGAGCTAAACTCTACTGCCTACAGGTGGGAGTGTGGGGAT CTGTACAGAATTGAGAAACAATAGGACAGGTGGGAGTGTGGGGATC TGAGCCAAACTGGAATGAACTT:GTGTGGAGACCCAGGCAGAGCA ATGGTATCAAAGGACAGTINS80bpAGGTTAGAATGAAGGATG GAGCAAAATTAAGGGAACAAGGGACAGGTGGGAGTGTGGGGATC AGCTGAGATGAGCTGGGGTGAGCCAGGTAAGAGTGTGGGAACC GGTTGTTAAAGAATGGTATCAAATGCAGGCAGAGTAGCTATAGGGCA GCCCTAGTAAGCGAGGCTCTINS98bpGGCAGAGCAGCTATAGGGC ACTGCCTACACTGGACTGTTCTGGGTGCCAGGACAGGTACAAGTTT	GCTAAACTAGGCTGGCTTAACCG:CCGGGTGAGCAAATACAAGGGAA GAGATGGGGTGAGATGGGGTGAG:AGAGAAGCTGAGGCAGGTAAGA ACTAGGCTGGCTTAACCG:CCGGGTGAGCAAATACAAGGGAA GAAAAGATGTTTTTAGTTTTINS215bpCCAGGCAGAGCAGCTATAG CCAGCCACAGTAATGACCCAGAGCAAATACAAGGGAACTGAT AGGGAACAAGGTTGAGAGCCAGCCAGGACAGGTGGGAGT TATGCTACGCTGTGTTGGGGTGAACTGATGGCAAATGGAAGGGCAG TGGGTGGGCTTCTCTGAGINS233bpCAGGGGCAGGTTAGAATGA GGATACGCAGAAGGAAGGCCACAGGAAGAATGCAGATCCAAAC
Mouse #2	CCCAGACAGAGAAAGCCAGACTGATGGCAAATGGAAGGGCA ACCCAGACAGAGAAAGCCA:CAAGACAACTAGAAGTGT ATGGCTGAGCTGAGATGGGGAGCCAGAACAGGTGGGA TGATTACTCTGAGGTAAGCAA:TACATACGGGTAAGCAGGGAC GAGAAAGAATAGAGACCTGCCCAGACTAAATGGCTACAGAG CCAGACTCATAAAGCTTGCTGACAACTAGAAGTGTGTGAAT AGGAAGGCCACAGCTGTACATAGGCAGGTAAGAGTGTGGG AATGTGGTTTAATGAATTTGAAINS87bpTAGAGGAACAGGGG TAAAAAGCATGGCTGAGCTGAGCCAAGACAACTAGA CCTACACTGGACTGTTCTGAGCAAATACAAGGGAACTGA	AGCTGTACAGAATTGAGAAAGAATGGGGATCCAGGTGCTGCAGCT ACAGTGCTTAGATCCGAGGTGA:CCAGGACAGGTGGGAGTGTGGG AGAGACCTGCAGTTGAGGCCAGCAGGTAAGCAGGGACAGGTGGA TGGGGTATGGATACGCAGAAGGAAGGGAAGTGATGGCAAATGG ACATTCTTGATCTACAACTCAATGAAGGATGTGCATCCTGGGTGAG TTGCTGAGCAAAATTAAGGGAA:ACGGGCAGGTTAGAATGAAG GGTGATTACTCTGAGGTAAGCAGAGGCAGGTAAGACTGTGGG GCCAGACTCATAAAGCTT:TGCCAGGACAGGTACAAGTTTAGT TAGGCTGGCTTAACCGAGATGAAGGATGGGCATCCCGGGTGAGC CTCAATGTGGTTTAATGAATTTGAA:CAGGTAAGGCAGGACTGGGGA
Mouse #3 Summary	GATCCAAGGTGAGTGTGAGAINS140bpTGGGGATCCAGGCAGTGT ATGAGCTGGGGTGAGCTCAGINS49bpTACAGGGAAGCTGAGGC AAGGTGAGTGTGAGAGGACAG:TGGGGAGCCAGGACAGGTGG CGCAGAAGGAAGGCCACAGCAGGGACAGGTGGAAGTGTG GAGGACAGGGGCTGGGGTATGGAGGAGCCAAGACAACTAGAAG TGAGCGCTTCTAAAATGCGCINS31bpTACAGGGAAGCTGAGGCAG TCTAAAAAGCATGGCTGAGCTG:TTAGGAGCAAGGACAGG CCTAGTAAGCGAGGCTCTAAINS38bpTACAGGGAAGCTGAGGCAGG AAGGCCAGACTCATAAAINS82bpTGCTGCAGCTACAGGTAAGCAG GATCCAAGGTGAGTGTGAGAINS140bpTGGGGATCCAGGCAGTGT	GGATACGCAGAAGGAAGGCCACACCAGGTGGGACTGTGGGGATC TGCGCTAAACTGAGGTGATTAAACCACAGAAGAGCAGGA GGATACGCAGAAGGAAGGCCACACCAGGTGGGACTGTGGGGATC ACGCAGAAGGAAGGCCACAG:AGGGTGCCAGGACAGGTAC AGCACAGCTGAGCTGAGATGGGGGAGTGTGGGGATCCAGGTAAG TGAGCTGAGCTAGGGTGAGCTGAGGGATGGGCATCCCGGGTGAG TTTTAGTTTTTATAGAGGACAGGGAAGCTATAGGG CAAAATTAAGGGAACAAGGTTGATGGGCATCCCGGGTGAGCAA AGCCAGACTCATAAAGCTTGCTINS191bpGGACAGGTGGAAGTG
Blunt	17%	28%
Nt. additions / insertions	33%	17%
1/2 bp microhom	30%	31%
>2 bp microhom	20%	24%

Figure S1. Sμ-Sγ1 junction recombination analysis in *$Zmynd8^{F/F}Cd19^{Cre/+}$* and control B cells.

Blunt junctions are indicated with "**:**", micro-homology in **bold**, mutations in *italics*, and nucleotide additions are underlined. Micro-homology at the junction was determined by identifying the longest region at the switch junction of perfect uninterrupted donor/acceptor identity. INS: insertion. Results from three mice per genotype are shown.

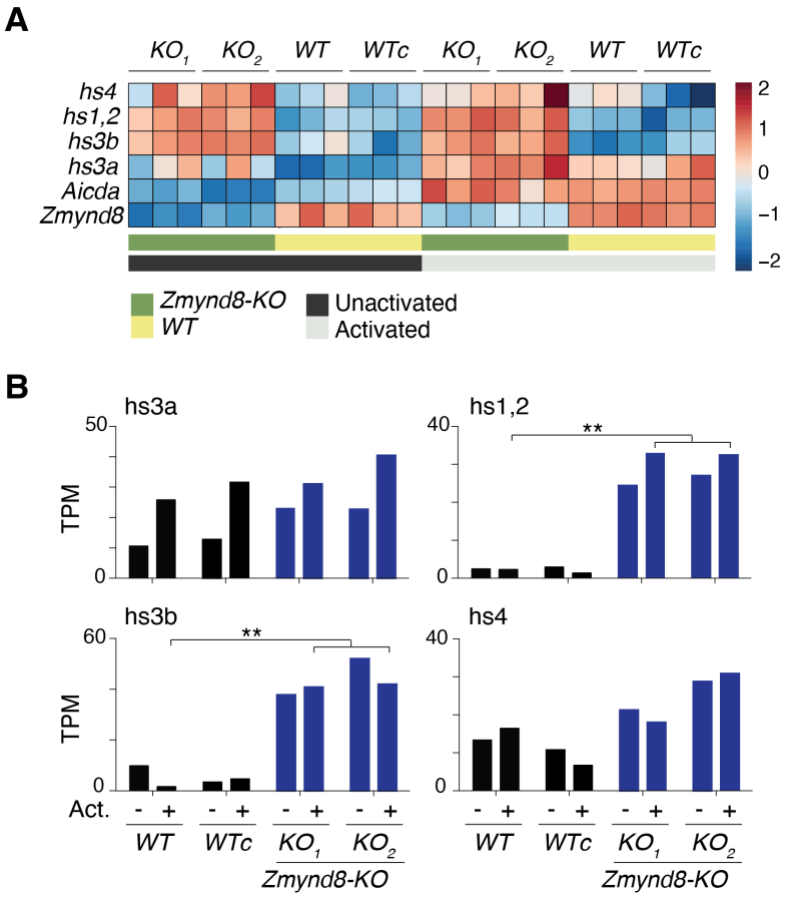

Figure S2. RNA-seq analysis of ZMYND8-deficient CH12 cells.
A. Heatmap showing Aicda, Zmynd8 and *Igh* 3'RR enhancers differential transcript expression as determined by RNA-seq in two control (*WT* bulk and a clonal derivative, *WTc*) and two *Zmynd8-KO* CH12 clones (KO_1 and KO_2). Expression counts are row-normalized by Z score. Data results from three replicates per sample. **B.** Graphs representing relative transcript levels of hs3a, hs3b, hs1,2 and hs4 in ZMYND8-deficient CH12 cells (blue) compared to controls (black). The adjusted p values were calculated with the Wald test and corrected for multiple testing with the Benjamini-Hochberg method.

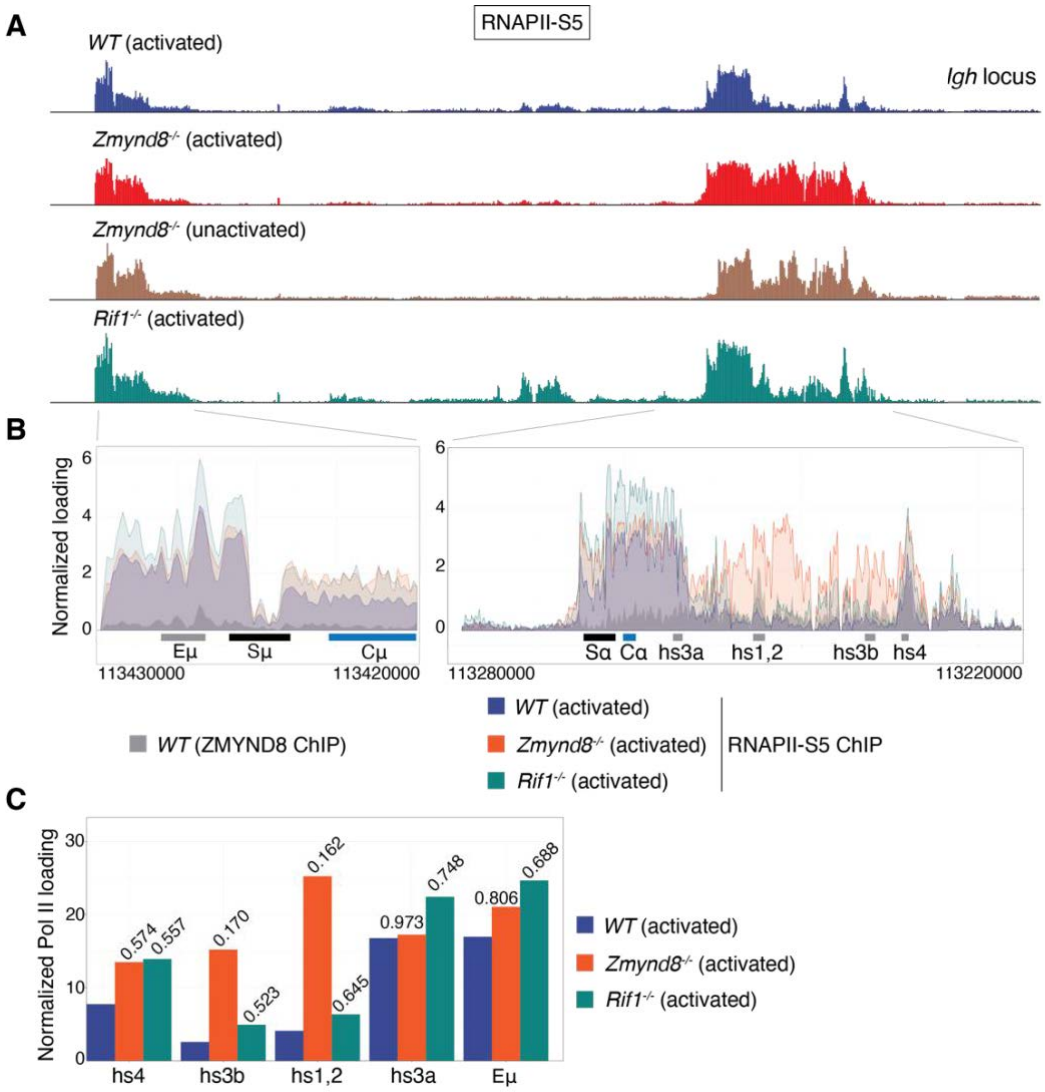

Figure S3. RNAPII-S5 is increased at the 3'RR in ZMYND8- but not in RIF1-deficient CH12 cells.

A. RNAPII-S5 ChIP-seq analysis at the Igh locus in *WT*, *Zmynd8*$^{-/-}$ and *Rif1*$^{-/-}$ CH12 cells. Cells were kept in complete RPMI (un-activated) or cultured in the presence of anti-CD40, TGFβ and IL4 (activated). **B.** Overlay of ChIP-seq tracks at Eμ and 3'RR enhancers. Regions were defined based on the location of ZMYND8 peaks from ZMYND8 ChIP-seq. **C.** RNAPII-S5 quantification in regions specified in B. Numbers represent the ration of RNAPII-S5 in *WT* versus *Zmynd8*$^{-/-}$ cells.

TABLES

Table 1. Genotyping primers

Primer name	Sequence (5'→3')	Reference
Zmynd8$^{F/F}$		
Prkcbp1_35576 (Fw)	GACCACAGCTCTTGCACAGG	*This study*
ZMYND8_R2 (Rev)	AAGAAAACCCTGAGACCACC	*This study*
MDV_p240 (Fw)	GTGCAAACGTGTTCAGTGG	*This study*
Pdap1$^{F/F}$		
MDV_p223 (Fw)	CTGGGGAAAAGGAGGCTCTG	*This study*
MDV_p313 (Fw)	GTCCTATGATTGGACCAAC	*This study*
MDV_p226 (Rev)	AACATACAGCCTGCCTTCCC	*This study*
Cd19$^{Cre/+}$		
p881	AACCAGTCAACACCCTTCC	*193*
p882	CCAGACTAGATACAGACCAG	*193*
p883	TCAGCTACACCAGAGACCGG	*193*
Zmynd8-KO (Exon 4)		
MDV_p100	CCTTCTGAACTGCCCTTTATG	*This study*
MDV_p101	GAAGGAAGAGACAGGTAGATTC	*This study*
Zmynd8-KO (Exon 5)		
MDV_p102	GGAAGAAAAGAGAAAGGAGAGG	*This study*
MDV_p103	GCAAATGACTGCTCAGAAAC	*This study*
Pdap1-KO (Exon 4)		
MDV_p228	ATGGGCTGTCGGTGGCTTG	*This study*
MDV_p229	GCTGTATGTGTAGCTGGGAG	*This study*

Table 2. gRNA sequences

gRNA	**Sequence (5'→3')**	**Reference**
Optimization loss-of-CSR screen		
g*Random*	GCGAGGTATTCGGCTCCGCG	*GeCKO (196)*
g*Random*	ATGTTGCAGTTCGGCTCGAT	*GeCKO (196)*
g*Random*	ACGTGTAAGGCGAACGCCTT	*GeCKO (196)*
g*Aicda-1*	CATGCCGTCCCTTGGCCCAG	*This study*
g*Aicda-2*	TCCTCTTCACCACGTAGCAG	*This study*
g*Aicda-3*	ACCTCCTGCTCACTGGACTT	*This study*
g*Aicda-4*	GTGACGCGGTAACACCGGCC	*This study*
g*Aicda-5*	CTATGACTGTGCCCGGCACG	*This study*
g*Aicda-6*	TTCACAGAAGTAGAGGCGCG	*This study*
g*Aicda-7*	TGCGGAGACTGCACCGCGCT	*This study*
g*Trp53bp1-1*	TGACGCGGGTGACGAGTGTA	*This study*
g*Trp53bp1-2*	CAGATGTTTATTATGTGGAT	*This study*
g*Trp53bp1-3*	GAGTGTACGGACTTCTCGAA	*This study*
g*Rif1-1*	CTCCAGGGGGCTGCGACCTG	*This study*
g*Rif1-2*	GGAAGACCCCTCGGTGCCTC	*This study*
g*Rif1-3*	AAGTCTCCAGAAGCGGCTCC	*This study*
g*Rif1-4*	GAAGACCCCTCGGTGCCTCC	*This study*
g*Rif1-5*	ACTCTTAATGATACCATTCA	*This study*
g*Rif1-6*	TGTGTGTACCAGGGCACTGT	*This study*

g*Rif1-7*	TAGGCGTCCGTCTGCTCTCC	*This study*
Loss-of-CSR screen (potential RIF1 interactors)		
g*Ahnak-1*	AACTCCCCTGCGGCCCGCAC	*This study*
g*Ahnak-2*	ACTCCCCTGCGGCCCGCACT	*This study*
g*Ahnak-3*	GGCTTGAAGTTGCACCGTAA	*This study*
g*Ahnak-4*	AACTCCCCTGCGGCCCGCAC	*This study*
g*Ahnak-5*	ACCACCCCAGTGCGGGCCGC	*This study*
g*Ahnak-6*	CTGGGCAATGGTCAGCCCGT	*This study*
g*Ccdc124-1*	GTTCTCGCCCTGGAACTTCT	*This study*
g*Ccdc124-2*	AGTCGGCAGCTGCCCGAGCA	*This study*
g*Ccdc124-3*	CCTTGGCGGGCGCCACACGT	*This study*
g*Ccdc124-4*	CACAGATCGAGGACTCGCTG	*This study*
g*Ccdc124-5*	GGACAAACACGTGATGCGGA	*This study*
g*Ccdc124-6*	CCCACGTGTGGCGCCCGCCA	*This study*
g*Dynll1-1*	GGAGAAGTACAACATCGAGA	*This study*
g*Dynll1-2*	CATGTCGGAAGAGATGCAAC	*This study*
g*Dynll1-3*	CATTTTTGATCACCGCCTTC	*This study*
g*Dynll1-4*	ACAAGAAGTACAACCCTACC	*This study*
g*Dynll1-5*	CAACATCGAGAAGGATATTG	*This study*
g*Dynll1-6*	TGCATTGTGGGCCGAAACTT	*This study*
g*Las1l-1*	TGGATCGCGTGTGGCGTGCG	*This study*
g*Las1l-2*	CCAAGAGACGACGATACCCC	*This study*
g*Las1l-3*	CGCTGAACCGAATTACAGTA	*This study*
g*Las1l-4*	CATCGATGCAGCCGGTACCT	*This study*
g*Las1l-5*	CGATGAGCTTGCATCTTACC	*This study*
g*Las1l-6*	GATGAACTTAGACTGCTCTA	*This study*
g*Mga-1*	CCTTCCGGCTGATTGTACTG	*This study*
g*Mga-2*	CTTCCGGCTGATTGTACTGT	*This study*
g*Mga-3*	TCCTCACCATCAAACCCGAA	*This study*
g*Mga-4*	AAGTTTCCATTCGGGTTTGA	*This study*
g*Mga-5*	TCCATTCGGGTTTGATGGTG	*This study*
g*Naa50-1*	TCACATCTCCCAGCTCGATC	*This study*
g*Naa50-2*	ACAGTTGAAGAGACTGAACC	*This study*
g*Naa50-3*	CTTGCACCTTACCGAAGACT	*This study*
g*Naa50-4*	CTAGTCTTCGGTAAGGTGCA	*This study*
g*Naa50-5*	ATTCTACAAGGATGTGCTAG	*This study*
g*Naa50-6*	CTGTAGGTGCAGTGTGCTGC	*This study*
g*Pyhin1-1*	AAAGGACACGGACAGGGAAA	*This study*
g*Pyhin1-2*	CATTTGTGGTGGACATAGGT	*This study*
g*Pyhin1-3*	GGTGTACTCCTCCCTGATTC	*This study*
g*Pyhin1-4*	ATACTGCTGGACGGTCTTTT	*This study*
g*Pyhin1-5*	AAGACCGTCCAGCAGTATTC	*This study*
g*Pyhin1-6*	CTAGAGGAACTCCTAGTGCC	*This study*
g*Rnf2-1*	AGTTCCCATGTTTTGCTTAA	*This study*
g*Rnf2-2*	GTGTTTACATCGGTTTTGCG	*This study*
g*Rnf2-3*	AGTGCATCAAAGTTCGGGTC	*This study*
g*Rnf2-4*	TGATGAGTATGAAGCGCATC	*This study*
g*Rnf2-5*	CTAGTCTTCGGTAAGGTGCA	*This study*
g*Rnf2-6*	CTGTAGGTGCAGTGTGCTGC	*This study*
g*Tex10-1*	CTATCCTCTTTAAGTTGTTC	*This study*
g*Tex10-2*	GCTGTAGCATTTTCTAACTT	*This study*
g*Tex10-3*	TTTGCCTGAACAACTTAAAG	*This study*
g*Tex10-4*	AATGTGTTCAGTTCGTATTT	*This study*

g*Tex10-5*	TGAGAAGTGACTCTCCGATT	*This study*
g*Tex10-6*	GCAACTCCATTTTTATCAAC	*This study*
g*Tlk2-1*	CAGTTAGCGCCACGGGGAGC	*This study*
g*Tlk2-2*	CACAGAGCAGTTAGCGCCAC	*This study*
g*Tlk2-3*	TAAAGTGGCCTAATCGCAAT	*This study*
g*Tlk2-4*	GAGCCTCGTTTACTGAGCAG	*This study*
g*Tlk2-5*	CGAGCAGAACAGCCTCTGTA	*This study*
g*Tlk2-6*	ATGCAGGACCGATTGCGATT	*This study*
g*Zmynd8-1*	AGAAAACGGCCCCGAAACGG	*This study*
g*Zmynd8-2*	AAGTCATTCCGTCCGTCCTG	*This study*
g*Zmynd8-3*	GCTGGGGAACCTCCGTTTCG	*This study*
g*Zmynd8-4*	TTGGGGCGAATGGCCATTGG	*This study*
g*Zmynd8-5*	TGGGGCTCGTGGATGAGTCT	*This study*
g*Zmynd8-6*	GACACTTAGCGTGATAAACC	*This study*
g*Zfp592-1*	TTCAAACGACTCTTGGCGGC	*This study*
g*Zfp592-2*	ACTTGTGTGCATGGATCCGC	*This study*
g*Zfp592-3*	CCATGGTGTTCCCCGGAACG	*This study*
g*Zranb2-1*	TCCGAGTCAGTGACGGGGAC	*This study*
g*Zranb2-2*	AATTTCCGAGTCAGTGACG	*This study*
g*Zranb2-3*	TCGACCACATCTGTTACAGC	*This study*
g*Zranb2-4*	GTGGAAATGTAAACTTTGCT	*This study*
g*Zranb2-5*	CATTGGCACTAAATAAGCCC	*This study*
g*Zranb2-6*	GCTTATTTAGTGCCAATGAT	*This study*
Loss-of-CSR screen (potential 53BP1 interactors)		
g*Pdap1-1*	AGTGAGGCAGTATACGAGCC	*This study*
g*Pdap1-2*	CATTGAGAACCCCAACCGCG	*This study*
g*Pdap1-3*	ACTGGATCTGGACGGGCCAA	*This study*
g*Mrpl14-1*	GATGACTCGGGTCCGAGTAG	*This study*
g*Mrpl14-2*	CACCGGGGTGGCCGATGGTA	*This study*
g*Mrpl14-3*	TTATAGACGTGGATGCACCG	*This study*
g*Eif4g2-1*	TAGAGTTGCGAGAGCACCAT	*This study*
g*Eif4g2-2*	TGCTCAGCTATGTCTGCGAT	*This study*
g*Eif4g2-3*	CATTTCTGAGCGTTTTGCCC	*This study*
g*Ddx18-1*	TGCAATACTAAGGGACCTTC	*This study*
g*Ddx18-2*	TTTGTGTCTTTTTAGCCCGT	*This study*
g*Ddx18-3*	GCCATCACCTGTCAGTCCTA	*This study*
g*Sec61a1-1*	CAGGTGGAATAGCGCTATGA	*This study*
g*Sec61a1-2*	ATAAGACCGGAGGTGACAAT	*This study*
g*Sec61a1-3*	ATCAAGTCGGCGCGTTACCG	*This study*
g*Ddx47-1*	GGACCAGGGCGAACAGTCGC	*This study*
g*Ddx47-2*	AAGTGGTCAATCAGTCGGCC	*This study*
g*Ddx47-3*	GACTCACCTATTACTATATG	*This study*
g*Chtop-1*	TCCCCCACGTAGTCCTCCTC	*This study*
g*Chtop-2*	CAGAGAGGCTTGCCCCGAGG	*This study*
g*Chtop-3*	GTCCCCCACGTAGTCCTCCT	*This study*
g*Pa2g4-1*	CCCAGGGGTGCTTCGATCTT	*This study*
g*Pa2g4-2*	GCCTTACGACTGGTCAAACC	*This study*
g*Pa2g4-3*	TCCAGGTTTGACCAGTCGTA	*This study*
g*Rsl1d1-1*	TTACTGATGCGAGAATTCGA	*This study*
g*Rsl1d1-2*	CAGCTGTAAAAGAGCGGTCG	*This study*
g*Rsl1d1-3*	TCCAGCTGTAAAAGAGCGGT	*This study*
g*Topbp1-1*	GAGATTATGATGACGACGTG	*This study*
g*Topbp1-2*	GACTCTACAGTTCGTGAAGA	*This study*

g*Topbp1-3*	CGTGCTTAGCCGTGAGTTGC	*This study*
g*Usp10-1*	GCAGAACTGCACACCGACGA	*This study*
g*Usp10-2*	TGCCGGCGGCCTTGGTCAGA	*This study*
g*Usp10-3*	GGCATGGCCGTTGACCAGGG	*This study*
g*Fus-1*	TCTGTTTTCTTGTAGCTACG	*This study*
g*Fus-2*	TCCAGCGGAAGTGACCGCGG	*This study*
g*Fus-3*	TTCTCCAGCGGAAGTGACCG	*This study*
g*G3bp1-1*	GCACACCTGTGCAACATGTC	*This study*
g*G3bp1-2*	CAGACCCTTACCTGTGACGC	*This study*
g*G3bp1-3*	AGCAACATACCTCAGGAGCA	*This study*
g*Fam76b-1*	CACGCAGCGCTACCCCTTCG	*This study*
g*Fam76b-2*	CCCTGCGAGAGCTCCTCGAA	*This study*
g*Fam76b-3*	CCCTTCGAGGAGCTCTCGCA	*This study*
g*Taf15-1*	CAAACGACTGATTCTTCGTA	*This study*
g*Taf15-2*	ATGACCGTCGTGATGTGAGT	*This study*
g*Taf15-3*	CATACCTACTCACATCACGA	*This study*
g*Snd1-1*	GGTGCGCCATAATTGTCCGA	*This study*
g*Snd1-2*	GGTCAATTGCAGTTTACACC	*This study*
g*Snd1-3*	ACGATGCCCCGCTGCACGGT	*This study*
g*Pabpc1l-1*	TACAATGAACATGGATCGCG	*This study*
g*Pabpc1l-2*	GGAAAGGATATTGCCGATCG	*This study*
g*Pabpc1l-3*	TGCCCATACCAGACTTCCGA	*This study*
g*Gtf2h2-1*	GTAGATGCGCCACCTGTATG	*This study*
g*Gtf2h2-2*	CTTAAGGAATTATAGAGCGA	*This study*
g*Gtf2h2-3*	GGTGGATGGATCGAGAACAA	*This study*
g*Ints10-1*	CACACTGGACCCGACCGGGA	*This study*
g*Ints10-2*	ATCCCACATCGTCGACCCCT	*This study*
g*Ints10-3*	GAATTTCCCGGACCAGCCGG	*This study*
g*Naca-1*	CACCTACCTGTCTCAGCCTG	*This study*
g*Naca-2*	TTATTCCCTTCAGAAATGCC	*This study*
g*Naca-3*	TGTAGCAGGGACGGTTTCTG	*This study*
g*Rpl22-1*	ATACTCCTAGAAAAAGCTTG	*This study*
g*Rpl22-2*	CAGAATGTCACTTACAAAAT	*This study*
g*Rpl22-3*	CTAGAAAAAGCTTGTGGCGA	*This study*
g*Sec61a2-1*	GCTACAGAAGGGTTACGGCT	*This study*
g*Sec61a2-2*	ATTTCCCCAATTGTAACGTC	*This study*
g*Sec61a2-3*	ATTAAGTCGGCACGGTATCG	*This study*
g*Ccdc124-1*	CTCACGGCTCAAAGGCGGCA	*This study*
g*Ccdc124-2*	CTTCCTGCAGGCGCCGGTTA	*This study*
g*Ccdc124-3*	CACAGATCGAGGACTCGCTG	*This study*
g*Mtdh1-1*	GAAGGAGCCCGCCCCGCAAA	*This study*
g*Mtdh1-2*	ACGGACTGTTGAAGTACCCG	*This study*
g*Mtdh1-3*	ACCCAGCCCGGGTACCGCTT	*This study*
g*Sub1-1*	GTCAGTGTTCGGGACTTCAA	*This study*
g*Sub1-2*	TTGCCTGTTTTAAGTTAAAG	*This study*
g*Sub1-3*	TGCTTCTTCACGGGCTTCTC	*This study*
g*Trip12-1*	AGTATAAACTCGTCCCAGAG	*This study*
g*Trip12-2*	CTACAACAGCAGATGACCGG	*This study*
g*Trip12-3*	TGTTACAGTCGGCGAGGCTC	*This study*
g*Abcf2-1*	GCCCGACAGCGACCCAGAAA	*This study*
g*Abcf2-2*	TCCATGCTGCTCTCGGCTAT	*This study*
g*Abcf2-3*	TCTGGGTCGCTGTCGGGCTT	*This study*
g*Rrbp1-1*	GCCACAGCGGATACACGCAC	*This study*

g*Rrbp1-2*	ACCAGTGCGTGTATCCGCTG	*This study*
g*Rrbp1-3*	TTCATACGACGTCTCCTTCA	*This study*
g*G3bp2-1*	GCTAGCTAGCCTTACGTCAC	*This study*
g*G3bp2-2*	TCATACCAAAATTCGTCATG	*This study*
g*G3bp2-3*	GGTCGCTGTTCCCGTACACG	*This study*
g*Larp4-1*	CCATCAATGCTCACCTCTAC	*This study*
g*Larp4-2*	TTTGATCTTACCGATTTTTT	*This study*
g*Larp4-3*	ATGTTGGCAACCGTCCATAT	*This study*
g*Hnrnpa2b-1*	GGCCCCATTCCATTGATGGC	*This study*
g*Hnrnpa2b-2*	ACCCTGCCATCAATGGAATG	*This study*
g*Hnrnpa2b-3*	AGCCACCAATAAAGAGCTTT	*This study*
g*Caprin-1*	CAGCATCCGGCAACCGGCAC	*This study*
g*Caprin-2*	GCGTAATCGACAAGAAACTT	*This study*
g*Caprin-3*	GCTTCTCAGCATCCGGCAAC	*This study*
g*Hnrnpa1-1*	AATTAACGCACCTCTCTGAC	*This study*
g*Hnrnpa1-2*	AATCTCTTTTAAGGTCGCAG	*This study*
g*Hnrnpa1-3*	TTTTAAGGTCGCAGTGGTTC	*This study*
CH12 KO clonal derivatives		
g*Zmynd8-1*	AGAAAACGGCCCCGAAACGG	*This study*
g*Zmynd8-2*	AAGTCATTCCGTCCGTCCTG	*This study*
gZmynd8 Nickase pair 1	GTCTTGGGGCGAATGGCCAT ATTAAAAAGAAAAAGAAACC	*This study*
gZmynd8 Nickase pair 2	GACACTTAGCGTGATAAACC GAGACTGACATCGGAGCCAG	*This study*
g*Trp53bp1* Nickase pair	CAGATGTTTATTATGTGGAT GAGTGTACGGACTTCTCGAA	*This study*
g*Rif1* Nickase pair	AAGTCTCCAGAAGCGGCTCC GAAGACCCCTCGGTGCCTCC	*This study*
g*Pdap1-1*	AGTGAGGCAGTATACGAGCC	*This study*
g*Pdap1-2*	CATTGAGAACCCCAACCGCG	*This study*
g*Pdap1-3*	ACTGGATCTGGACGGGCCAA	*This study*
g*Pdap1-4*	GAAGGTCACGCAACTGGATC	*This study*
g*Pdap1-5*	TTGTCTGTGCCACGCGGTTG	*This study*
g*Pdap1-6*	GTTGTCTGTGCCACGCGGTT	*This study*

Table 3. SHM primers

Name	**Sequence (5'→3')**	**Reference**
J_H4		
VHA (Fw)	ARGCCTGGGRCTTCAGTGAAG	*201*
VHE (Fw)	GTGGAGTCTGGGGGAGGCTTA	*201*
JH4_intron (Rev)	CTCCACCAGACCTCTCTAGACAGC	*201*
J_K5		
VK (Fw)	GGCTGCAGSTTCAGTGGCAGTGGRTCWG GRAC	*24*
JK5_PR (Rev)	AGCGAATTCAACTTAGGAGACAAAAGAGA GAAC	*24*

Table 4. RT-qPCR primers

Name	Sequence (5'→3')	Reference
Aicda		
AID-F (Fw)	GAAAGTCACGCTGGAGACCG	*203*
AID-R (Rev)	TCTCATGCCGTCCCTTGG	*203*
pre-spliced GLT γ1		
MDV_g1-3b (Fw)	CAGGATCAATCCCAGCATTGGG	*This study*
MDV_g1-1 (Rev)	CTGTGCTTGGATCACCACACTTCC	*202*
pre-spliced GLT γ3		
MDV_g3-3 (Fw)	GTGGAACTCTAAGGTTTAGGAGTCAA	*202*
MDV_g3-1 (Rev)	CTGTGGCTGCTCAACTTGGTACCTT	*This study*
pre-spliced GLT γ2b		
MDV_g2b-3 (Fw)	GTTGACCTGACCTAGAGACTGGTGGAC	*202*
MDV_g2b-1 (Rev)	CTTTCTTTCAGCTTCATTCATGGAAC	*202*
post-spliced GLT γ1		
Ig1 (Fw)	GGCCCTTCCAGATCTTTGAG	*53*
Cg1R (Rev)	GGATCCAGAGTTCCAGGTCACT	*53*
post-spliced GLT γ3		
Ig3F (Fw)	TGGGCAAGTGGATCTGAACA	*53*
Cg3R (Rev)	CTCAGGGAAGTAGCCTTTGACA	*53*
post-spliced GLT γ2b		
MDV_p243 (Fw)	CACTGGGCCTTTCCAGAACTA	*53*
MDV_p244 (Rev)	CACTGAGCTGCTCATAGTGTAGAGTC	*53*
post-spliced GLT μ		
ImF (Fw)	CTCTGGCCCTGCTTATTGTTG	*53*
CmR (Rev)	GAAGACATTTGGGAAGGACTGACT	*53*
post-spliced GLT α		
IaF (Fw)	CCTGGCTGTTCCCCTATGAA	*53*
CaR (Rev)	GAGCTGGTGGGAGTGTCAGTG	*53*
3'RR hs1,2		
MDV_p324 (Fw)	CATTGAGCTCCGGCTCTAAC	*This study*
MDV_p325 (Rev)	CAAGAGGACATGACAGGAGATG	*This study*
3'RR 5' hs3b		
MDV_p314 (Fw)	CATTGAGCTCCGGCTCTAAC	*This study*
MDV_p315 (Rev)	CCCCTGTAGGGATCCTCCTAAT	*This study*
3'RR 3' hs3b		
MDV_p316 (Fw)	CATCCAGAGTCAAGGGGTGTC	*This study*
MDV_p317 (Rev)	CTAGAACCACATGCTATCTAAGGGA	*This study*
Housekeeping		
Ubc_Fw	GCCCAGTGTTACCACCAAGA	305
Ubc_Rev	CCCATCACACCCAAGAACA	305
Gapdh_Fw	TGTGTCCGTCGTGGATCTGA	305

Gapdh_Rev	TTGCTGTTGAAGTCGCAGGAG	305

Table 5. MutPE-seq primers

Name	Sequence (5'→3')	Reference
First PCR		
QW_506 (Fw)	TCTACACTCTTTCCCTACACGACGCTCTTC CGATCTTCTCTGAGTGCTTCTAAAATGCG	93
QW_507 (Rev)	GTGACTGGAGTTCAGACGTGTGCTCTTCC GATCTTCACCCCAACACAGCGTAGC	93
Second PCR		
QW_501 (Fw)	AATGATACGGCGACCACCGAGATCTACAC TCTTTCCCTACACGAC	93
QW_OutIndex (Rev)	CAAGCAGAAGACGGCATACGAGAT*6ntInde x*GTGACTGGAGTTCAGACGTGTG	93

Table 6. Switch junction analysis primers

Name	Sequence (5'→3')	Reference
First PCR		
5m3 (Fw)	AATGGATACCTCAGTGGTTTTTAATGGTG GGTTTAATATAG	*96*
gamma1R (Rev)	TGCCAATTAGCTCCTGCTCTTCTGTGG	*96*
Second PCR		
5m3-1 (Fw)	GGCTAAGAAGGCAATCCTGGGATTCTGG	*96*
gamma1R-1 (Rev)	CTCTTACCTGCCTAGCTTCTCTGTAGCC	*96*

REFERENCES

1. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **124**, 783–801 (2006).
2. Flajnik, M. F. & Masanori, K. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet.* **11**, 47–59 (2010).
3. Farber, D. L., Netea, M. G., Radbruch, A., Rajewsky, K. & Zinkernagel, R. M. Immunological memory: lessons from the past and a look to the future. *Nat. Rev. Immunol.* **16**, 124–28 (2016).
4. Chaplin, D. D. Overview of the Immune Response. *J Allergy Clin Immunol* **125**, S3-23 (2010).
5. Schroeder, H. W. J. & Cavacini, L. Structure and Function of Immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41–S52 (2010).
6. Janeway, C. J., Travers, P., Walport, M. & Al., E. *Immnobiology: The Immune System in health and disease. 5th edition.* (2001).
7. Bassing, C. H., Swat, W. & Alt, F. W. The mechanism and regulation of chromosomal V(D)J recombination. *Cell* **109**, 45–55 (2002).
8. Briney, B., Inderbitzin, A., Joyce, C. & Burton, D. R. Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* **566**, 393–397 (2019).
9. Porter, R. R. The hydrolysis of rabbit y-globulin and antibodies with crystalline papain. *Biochem. J.* **73**, 119–126 (1959).
10. Xu, Z., Zan, H., Pone, E. J., Mai, T. & Casali, P. Immunoglobulin class switch DNA recombination: induction, targeting and beyond. *Nat. Rev. Immunol.* **12**, 517–31 (2013).
11. Lu, L. L., Suscovich, T. J., Forture, S. M. & Alter, G. Beyond binding: antibody effector functions in infectious diseases. *Nat. Rev. Immunol.* **18**, 46–61 (2018).
12. Pelanda, R. & Torres, R. M. Central B-Cell tolerance: Where selection begins. *Cold Spring Harb. Perspect. Biol.* **4**, 1–15 (2012).
13. Johnston, C. M., Wood, A. L., Bolland, D. J. & Corcoran, A. E. Complete Sequence Assembly and Characterization of the C57BL/6 Mouse Ig Heavy Chain V Region. *J. Immunol.* **176**, 4221–4234 (2006).
14. Ye, J. The immunoglobulin IGHD gene locus in C57BL/6 mice. *Immunogenetics* **56**, 399–404 (2004).
15. Vettermann, C. & Schlissel, M. S. Allelic exclusion of immunoglobulin genes: Models and mechanisms. *Immunol. Rev.* **237**, 22–42 (2010).
16. Marquet, M. *et al.* The E Enhancer Region Influences H Chain Expression and B Cell Fate without Impacting IgVH Repertoire and Immune Response In Vivo. *J. Immunol.* **193**, 1171–1183 (2014).
17. Banerji, J., Olson, L. & Schaffner, W. A Lymphocyte-Specific Cellular Enhancer Is Located Downstream of the Joining Region in Immunoglobulin Heavy Chain Genes. *Cell* **33**, 729–740 (1983).

18. Perlot, T., Alt, F. W., Bassing, C. H., Suh, H. & Pinaud, E. Elucidation of IgH intronic enhancer functions via germ-line deletion. *Proc. Natl. Acad. Sci.* **102**, 14362–14367 (2005).
19. Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* **33**, 717–728 (1983).
20. Matthews, A. J., Zheng, S., DiMenna, L. J. & Chaudhuri, J. *Regulation of Immunoglobulin Class-Switch Recombination: Choreography of Noncoding Transcription, Targeted DNA Deamination, and Long-Range DNA repair*. *Advances in Immunology* vol. 122 (Elsevier Inc., 2014).
21. Geisberger, R., Lamers, M. & Achatz, G. The riddle of the dual expression of IgM and IgD. *Immunology* **118**, 429–437 (2006).
22. Vincent-fabert, C., Fiancette, R., Pinaud, E., Cogne, N. & Cogne, M. Brief report Genomic deletion of the whole IgH 3J regulatory region (hs3a , hs1 , 2 , hs3b , and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. *Hematology* **116**, 1895–1898 (2010).
23. Saintamand, A. *et al.* The IgH 3' regulatory region governs μ chain transcription in mature B lymphocytes and the B cell fate. *Oncotarget* **6**, 4845–4852 (2015).
24. Rouaud, P. *et al.* The IgH 3′ regulatory region controls somatic hypermutation in germinal center B cells. *J. Exp. Med.* **210**, 1501–1507 (2013).
25. Pinaud, E. *et al.* Localization of the 3′ IgH locus elements that effect long-distance regulation of class switch recombination. *Immunity* **15**, 187–199 (2001).
26. Manis, J. P. *et al.* Class Switching in B Cells Lacking 3′ Immunoglobulin Heavy Chain Enhancers. *J. Exp. Med.* **188**, 1421–1431 (1998).
27. Cogné, M. *et al.* A class switch control region at the 3′ end of the immunoglobulin heavy chain locus. *Cell* **77**, 737–747 (1994).
28. Predeus, A. V. *et al.* Targeted Chromatin Profiling Reveals Novel Enhancers in Ig H and Ig L Chain Loci. *J. Immunol.* **192**, 1064–1070 (2014).
29. Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
30. Medvedovic, J. *et al.* Flexible long-range loops in the V H gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. *Immunity* **39**, 229–244 (2013).
31. Amoretti-Villa, R., Rogier, M., Robert, I., Heyer, V. & Reina-San-Martin, B. A novel regulatory region controls IgH locus transcription and switch recombination to a subset of isotypes. *Nat. Cell. Mol. Immunol.* (2019).
32. Woloschak, G. & Krco, C. Regulation of kappa/lambda immunoglobulin light chain expression in normal murine lymphocytes. *Mol. Immunol.* **24**, 751–7 (1987).
33. Collins, A. M. & Watson, C. T. Immunoglobulin light chain gene rearrangements, receptor editing and the development of a self-tolerant antibody repertoire. *Front. Immunol.* **9**, 1–12 (2018).

34. Montecino-Rodriguez, E. & Dorshkind, K. B-1 B Cell Development in the Fetus and Adult. *Immunity* **36**, 13–21 (2012).
35. Seita, J. & Weissman, I. . Hematopoietic Stem Cell: Self-renewal versus Differentiation. *Wiley Interdiscip Rev.Syst. Biol.Med.* **2**, 640–53 (2010).
36. Pieper, K; Grimbacher, B; Eibel, H. B-cell biology and development. *J Allergy Clin Immunol* **131**, 959–71 (2013).
37. Hardy, R. R. & Hayakawa, K. B cell deveopment pathways. *Annu. Rev. Immunol.* **19**, 595–621 (2001).
38. Meffre, E., Casellas, R. & Nussenzweig, M. C. Antibody regulation of B cell development. *Nat. Immunol.* **1**, 379–385 (2000).
39. Allman, D. *et al.* Resolution of Three Nonproliferative Immature Splenic B Cell Subsets Reveals Multiple Selection Points During Peripheral B Cell Maturation. *J. Immunol.* **167**, 6834–6840 (2001).
40. Srivastava, B., Quinn, W. J., Hazard, K., Erikson, J. & Allman, D. Characterization of marginal zone B cell precursors. *J. Exp. Med.* **202**, 1225–1234 (2005).
41. Cerutti, A., Cols, M. & Puga, I. Marginal zone B cells: virtues of innatelike antibody-producing lymphocytes. *Nat. Rev. Immunol.* **13**, 118–32 (2013).
42. Gutzeit, C., Chen, K. & Cerutti, A. The enigmatic function of IgD: some answers at last. *Eur. J. Immunol.* **48**, 1101–113 (2018).
43. Victora, G. D. & Nussenzweig, M. C. Germinal Centers. *Annu. Rev. Immunol.* **30**, 429–457 (2012).
44. Dörner, T. & Radbruch, A. Selecting B cells and plasma cells to memory. *J. Exp. Med.* **201**, 497–499 (2005).
45. Rajewsky, K., Förster, I. & Cumano, A. Evolutionary and somatic selection of the antibody repertoire in the mouse. *Science (80-.).* **238**, 1088–94 (1987).
46. Rogozin, I. B. & Kolchanov, N. A. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim. Biophys. Acta* **1171**, 11–8 (1992).
47. Pham, P., Bransteitter, R., Petruska, J. & Goodman, M. F. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* **424**, 103–7 (2003).
48. Rogozin, I. B., Pavlov, Y. I., Bebenek, K., Matsuda, T. & Kunkel, T. A. Somatic mutation hotspots correlate with DNA polymerase eta error spectrum. *Nat. Immunol.* **2**, 530–6 (2001).
49. Rogozin, I. B. & Diaz, M. Cutting Edge: DGYW/WRCH Is a Better Predictor of Mutability at G:C Bases in Ig Hypermutation Than the Widely Accepted RGYW/WRCY Motif and Probably Reflects a Two-Step Activation-Induced Cytidine Deaminase-Triggered Process. *J. Immunol.* **172**, 3382–3384 (2004).
50. Pham, P., Calabrese, P., Park, S. J. & Goodman, M. F. Analysis of a single-stranded DNA-scanning process in which activation-induced deoxycytidine deaminase (AID)

deaminates C to U haphazardly and inefficiently to ensure mutational diversity. *J. Biol. Chem.* **286**, 24931–24942 (2011).

51. Ramiro, A. R., Stavropoulos, P., Jankovic, M. & Nussenzweig, M. C. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat. Immunol.* **4**, 452–456 (2003).
52. Dickerson, S. K., Market, E., Besmer, E. & Papavasiliou, F. N. AID Mediates Hypermutation by Deaminating Single Stranded DNA. *J. Exp. Med.* **197**, 1291–1296 (2003).
53. Muramatsu, M. *et al.* Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563 (2000).
54. Odegard, V. H. & Schatz, D. G. Targeting of somatic hypermutation. *Nat. Rev. Immunol.* **6**, 573–583 (2006).
55. Papavasiliou, F. N. & Schatz, D. G. Somatic hypermutation of immunoglobulin genes: Merging mechanisms for genetic diversity. *Cell* **109**, 35–44 (2002).
56. Teng, G. & Papavasiliou, F. N. Immunoglobulin Somatic Hypermutation. *Annu. Rev. Genet.* **41**, 107–120 (2007).
57. Methot, S. P. & Di Noia, J. M. *Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination*. *Advances in Immunology* vol. 133 (Elsevier Inc., 2017).
58. Yewdell, W. T. & Chaudhuri, J. A transcriptional serenAID: The role of noncoding RNAs in class switch recombination. *Int. Immunol.* **29**, 183–196 (2017).
59. Manis, J. P., Tian, M. & Alt, F. W. Mechanism and control of class-switch recombination. *Trends Immunol.* **23**, 31–39 (2002).
60. Stavnezer, J. *et al.* Immunoglobulin heavy-chain switching may be directed by prior induction of transcripts from constant-region genes. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 7704–7708 (1988).
61. Rothman, P., Lutzker, S., Cook, W., Coffman, R. & Alt, F. W. Mitogen plus interleukin 4 induction of C epsilon transcripts in B lymphoid cells. *J Exp Med* **168**, 2385 (1988).
62. Severinson, E., Fernandez, C. & Stavnezer, J. Induction of germ-line immunoglobulin heavy chain transcripts by mitogens and interleukins prior to switch recombination. *Eur. J. Immunol.* **20**, 1079–1084 (1990).
63. Mao, C. & Stavnezer, J. Differential Regulation of Mouse Germline Ig γ1 and ε Promoters by IL-4 and CD40. *J. Immunol.* **167**, 1522–1534 (2001).
64. Stavnezer, J. & Kang, J. The Surprising Discovery That TGFβ Specifically Induces the IgA Class Switch. *J. Immunol.* **182**, 5–7 (2009).
65. Park, S. R., Seo, G. Y., Choi, A. J., Stavnezer, J. & Kim, P. H. Analysis of transforming growth factor-β1-induced Ig germ-line γ2b transcription and its implication for IgA isotype switching. *Eur. J. Immunol.* **35**, 946–956 (2005).
66. Lebman, D. a, Nomura, D. Y., Coffman, R. L. & Lee, F. D. Molecular characterization of germ-line immunoglobulin A transcripts produced during transforming growth factor type

beta-induced isotype switching. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 3962–6 (1990).

67. Kaminski, D. A. & Stavnezer, J. Stimuli that enhance IgA class switching increase histone 3 acetylation at Sα, but poorly stimulate sequential switching from IgG2b. *Eur. J. Immunol.* **37**, 240–251 (2007).
68. McIntyre, T. M., Kehry, M. R. & Snapper, C. M. Novel in vitro model for high-rate IgA class switching. *J. Immunol.* **154**, 3156–61 (1995).
69. Snapper, C. & Paul, W. Interferon-gamma and B cell stimulatory factor-1 reciprocally regulate Ig isotype production. *Science (80-.).* **236**, 944–7 (1987).
70. Stavnezer, J. Mechanism and regulation of class switch recombination. *Annu. Rev. Immunol.* **26**, 261–92 (2008).
71. Stavnezer, J. & Schrader, C. E. IgH Chain Class Switch Recombination: Mechanism and Regulation. *J. Immunol.* **193**, 5370–5378 (2014).
72. Chaudhuri, J. & Alt, F. W. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat. Rev. Immunol.* **4**, 541–552 (2004).
73. Basu, U. *et al.* The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* **144**, 353–363 (2011).
74. Conticello, S. G. *et al.* Interaction between Antibody-Diversification Enzyme AID and Spliceosome-Associated Factor CTNNBL1. *Mol. Cell* **31**, 474–484 (2008).
75. Zheng, S. *et al.* Non-coding RNA generated following lariat-debranching mediates targeting of AID to DNA Simin. *Cell* **176**, 762–73 (2015).
76. Pavri, R. *et al.* Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* **143**, 122–133 (2010).
77. Voung, B. Q. *et al.* A DNA break– and phosphorylation-dependent positive feedback loop promotes immunoglobulin class-switch recombination. *Nat. Immunol.* **14**, 1183–89 (2013).
78. Cheng, H. L. *et al.* Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 2717–2722 (2009).
79. Basu, U. *et al.* The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature* **438**, 508–11 (2005).
80. McBride, K. M. *et al.* Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J. Exp. Med.* **205**, 2585–2594 (2008).
81. Wang, L., Wuerffel, R., Feldman, S., Khamlichi, A. A. & Kenter, A. L. S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *J. Exp. Med.* **206**, 1817–30 (2009).
82. Vaidyanathan, B. & Chaudhuri, J. Epigenetic Codes Programing Class Switch Recombination. *Front. Immunol.* **6**, 405 (2015).
83. Zan, H. & Casali, P. Epigenetics of peripheral B-cell differentiation and the antibody response. *Front. Immunol.* **6**, (2015).
84. Wang, Q. *et al.* Epigenetic targeting of activation-induced cytidine deaminase. *Proc.*

Natl. Acad. Sci. U. S. A. **111**, 18667–18672 (2014).

85. Daniel, J. A. *et al.* PTIP Promotes Chromatin Changes Critical for Immunoglobulin Class Switch Recombination. *Science (80-.).* **329**, 917–23 (2010).
86. Schwab, K. R., Patel, S. R. & Dressler, G. R. Role of PTIP in Class Switch Recombination and Long-Range Chromatin Interactions at the Immunoglobulin Heavy Chain Locus. *Mol. Cell. Biol.* **31**, 1503–1511 (2011).
87. Xu, Z. *et al.* 14-3-3 adaptor proteins recruit AID to 5′-AGCT-3′-rich switch regions for class switch recombination. *Nat. Struct. Mol. Biol.* **17**, 1124–1135 (2010).
88. Jeevan-Raj, B. P. *et al.* Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J. Exp. Med.* **208**, 1649–1660 (2011).
89. Brandsma, I. & Gent, D. C. Pathway choice in DNA double strand break repair: observations of a balancing act. *Genome Integr.* **3**, 9 (2012).
90. Goodarzi, A. a. & Jeggo, P. a. *The Repair and Signaling Responses to DNA Double-Strand Breaks*. *Advances in Genetics* vol. 82 (Elsevier Inc., 2013).
91. Jasin, M. & Rothstein, R. Repair of strand breaks by homologous recombination. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).
92. He, M., Cortizas, E. M., Verdun, R. E. & Severinson, E. Cyclin-Dependent Kinases Regulate Ig Class Switching by Controlling Access of AID to the Switch Region. *J. Immunol.* **194**, 4231–4239 (2015).
93. Wang, Q. *et al.* The cell cycle restricts activation-induced cytidine deaminase activity to early G1. *J. Exp. Med.* **214**, 49–58 (2017).
94. Boboila, C., Alt, F. W. & Schwer, B. *Classical and Alternative End-Joining Pathways for Repair of Lymphocyte-Specific and General DNA Double-Strand Breaks*. *Advances in Immunology* vol. 116 (Elsevier Inc., 2012).
95. Deriano, L. & Roth, D. B. Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage. *Annu. Rev. Genet.* **47**, 433–455 (2013).
96. Di Virgilio, M. *et al.* Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. *Science* **339**, 711–5 (2013).
97. Chapman, J. R. *et al.* RIF1 Is Essential for 53BP1-Dependent Nonhomologous End Joining and Suppression of DNA Double-Strand Break Resection. *Mol. Cell* **49**, 858–871 (2013).
98. Bothmer, A. *et al.* Regulation of DNA End Joining, Resection, and Immunoglobulin Class Switch Recombination by 53BP1. *Mol. Cell* **42**, 319–329 (2011).
99. Callen, E. *et al.* 53BP1 mediates productive and mutagenic DNA repair through distinct phosphoprotein interactions. *Cell* **153**, 1266–1280 (2013).
100. Zimmermann, M., Lottersberger, F. & Buonomo, S. B. 53BP1 Regulates DSB Repair Using Rif1 to Control 5' End Resection. *Science (80-.).* **339**, (2013).
101. Reina-San-Martin, B., Chen, J., Nussenzweig, A. & Nussenzweig, M. C. Enhanced intra-switch region recombination during immunoglobulin class switch recombination in

53BP1-/- B cells. *Eur. J. Immunol.* **37**, 235–239 (2007).

102. Boboila, C. *et al.* Alternative end-joining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase. *J. Exp. Med.* **207**, 417–427 (2010).
103. Hodgkin, P. D., Lee, J. H. & Lyons, a B. B cell differentiation and isotype switching is related to division cycle number. *J. Exp. Med.* **184**, 277–281 (1996).
104. Deenick, E. K., Hasbold, J. & Hodgkin, P. D. Switching to IgG3, IgG2b, and IgA is division linked and independent, revealing a stochastic framework for describing differentiation. *J. Immunol.* **163**, 4707–14 (1999).
105. Hasbold, J., Sui-Yin Hong, J., Kehry, M. R. & Hod. Integrating Signals from IFN-γ and IL-4 by B Cells: Positive and Negative Effects on CD40 Ligand-Induced Proliferation, Survival, and Division-Linked Isotype Switching to IgG1, IgE and IgG2a. *J. Immunol.* **163**, 4175–81 (1999).
106. Hasbold, J., Lyons, A. B., Kehry, M. R. & Hodgkin, P. D. Cell division number regulates IgG1 and IgE switching of B cells following stimulation by CD40 ligand and IL-4. *Eur. J. Immunol.* **28**, 1040–1051 (1998).
107. Hasham, M. G. *et al.* Activation-Induced Cytidine Deaminase-Initiated Off-Target DNA Breaks Are Detected and Resolved during S Phase. *J. Immunol.* **189**, 2374–2382 (2012).
108. Maul, R. W. *et al.* Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions Robert. *Nat. Immunol.* **12**, 70–76 (2011).
109. Sharbeen, G., Yee, C. W. Y., Smith, A. L. & Jolly, C. J. Ectopic restriction of DNA repair reveals that UNG2 excises AID-induced uracils predominantly or exclusively during G1 phase. *J. Exp. Med.* **209**, 965–974 (2012).
110. Petersen, S. *et al.* AID is required to initiate Nbs1/γ-H2AX focus formation and mutations at sites of class switching. *Nature* **414**, 660–665 (2001).
111. Rush, J. S., Liu, M., Odegard, V. H., Unniraman, S. & Schatz, D. G. Expression of activation-induced cytidine deaminase is regulated by cell division, providing a mechanistic basis for division-linked class switch recombination. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13242–7 (2005).
112. Wiedemann, E. M., Peycheva, M. & Pavri, R. DNA Replication Origins in Immunoglobulin Switch Regions Regulate Class Switch Recombination in an R-Loop-Dependent Manner. *Cell Rep.* **17**, 2927–2942 (2016).
113. Norio, P. *et al.* Progressive activation of DNA replication initiation in large domains of the immunoglobulin heavy chain locus during B cell development. *Mol. Cell* **20**, 575–587 (2005).
114. Ju, Z. *et al.* Evidence for physical interaction between the immunoglobulin heavy chain variable region and the 3′ regulatory region. *J. Biol. Chem.* **282**, 35169–35178 (2007).
115. Wuerffel, R. *et al.* S-S Synapsis during Class Switch Recombination Is Promoted by Distantly Located Transcriptional Elements and Activation-Induced Deaminase. *Immunity* **27**, 711–722 (2007).

116. Gostissa, M. *et al.* IgH class switching exploits a general property of two DNA breaks to be joined in cis over long chromosomal distances. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 2644–2649 (2014).
117. Dong, J. *et al.* Orientation-Specific Joining of AID-initiated DNA Breaks Promotes Antibody Class Switching. *Nature* **525**, 134–139 (2015).
118. Feldman, S. *et al.* 53BP1 Contributes to Igh Locus Chromatin Topology during Class Switch Recombination . *J. Immunol.* **198**, 2434–2444 (2017).
119. Rocha, P. P. *et al.* A Damage-Independent Role for 53BP1 that Impacts Break Order and Igh Architecture during Class Switch Recombination. *Cell Rep.* **16**, 48–55 (2016).
120. Giannini, S. L., Singh, M., Calvo, C. F., Ding, G. & Birshtein, B. K. DNA regions flanking the mouse Ig 3'α enhancer are differentially methylated and DNAse I hypersensitive during B cell differentiation. *J. Immunol.* **150**, 1772–1780 (1993).
121. Lieberson, R., Giannini, S. L., Birshtein, B. K. & Eckhardt, L. A. An enhancer at the 3′ end of the mouse immunoglobulin heavy chain locus. *Nucleic Acids Res.* **19**, 933–937 (1991).
122. Madisen, L. & Groudine, M. Identification of a locus control region in the immunoglobulin heavy-chain locus that deregulates c-myc expression in plasmacytoma and Burkitt's lymphoma cells. *Genes Dev.* **8**, 2212–2226 (1994).
123. Matthias, P. & Baltimore, D. The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the alpha constant region. *Mol. Cell. Biol.* **13**, 1547–1553 (1993).
124. Michaelson, J. S., Giannini, S. L. & Birshtein, B. K. Identification of 3α-hs4, a novel Ig heavy chain enhancer element regulated at multiple stages of B cell differentiation. *Nucleic Acids Res.* **23**, 975–981 (1995).
125. Pettersson, S., Cook, G., Brüggemann, M., Williams, G. & Neuberger, M. S. A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. *Nature* **344**, 165–8 (1990).
126. Birshtein, B. K. Epigenetic regulation of individual modules of the immunoglobulin heavy chain locus 3' regulatory region. *Front. Immunol.* **5**, 1–9 (2014).
127. Pinaud, E. *et al.* The IgH locus 3' regulatory region: pulling the strings from behind. *Adv. Immunol.* **110**, 27–70 (2011).
128. Le Noir, S. *et al.* Functional anatomy of the immunoglobulin heavy chain 3' super-enhancer needs not only core enhancer elements but also their unique DNA context. *Nucleic Acids Res.* **45**, 5829–5837 (2017).
129. Garot, A. *et al.* Sequential activation and distinct functions for distal and proximal modules within the IgH 3′ regulatory region. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1618–1623 (2016).
130. Manis, J., Michaelson, J., Birshtein, B. & Alt, F. Elucidation of a downstream boundary of the 3' IgH regulatory region. *Mol. Immunol.* **39**, 753–60 (2003).
131. Rouaud, P. *et al.* The IgH 3' regulatory region controls somatic hypermutation in

germinal center B cells. *J. Exp. Med.* **210**, 1501–1507 (2013).

132. Vincent-Fabert, C. *et al.* Ig Synthesis and Class Switching Do Not Require the Presence of the hs4 Enhancer in the 3′ IgH Regulatory Region. *J. Immunol.* **182**, 6926–6932 (2009).
133. Bébin, A.-G. *et al.* In Vivo Redundant Function of the 3′ IgH Regulatory Element HS3b in the Mouse . *J. Immunol.* **184**, 3710–3717 (2010).
134. Saintamand, A. *et al.* Elucidation of IgH 3′ region regulatory role during class switch recombination via germline deletion. *Nat. Commun.* **6**, (2015).
135. Saintamand, A. *et al.* Deciphering the importance of the palindromic architecture of the immunoglobulin heavy-chain 3' regulatory region. *Nat. Commun.* **7**, (2016).
136. Khamlichi, A., Pinaud, E., Decourt, C., Chauveau, C. & Cogné, M. The 3' IgH regulatory region: a complex structure in a search for a function. *Adv. Immunol.* **75**, 317–45 (2000).
137. Escribano-Díaz, C. *et al.* A Cell Cycle-Dependent Regulatory Circuit Composed of 53BP1-RIF1 and BRCA1-CtIP Controls DNA Repair Pathway Choice. *Mol. Cell* **49**, 872–883 (2013).
138. Feng, L., Fong, K. W., Wang, J., Wang, W. & Chen, J. RIF1 counteracts BRCA1-mediated end resection during DNA repair. *J. Biol. Chem.* **288**, 11135–11143 (2013).
139. Gherzraoui, H. *et al.* 53BP1 cooperation with the REV7-shieldin complex underpins DNA structure-specific NHEJ. *Nature* **560**, 122–7 (2018).
140. Barazas, M. *et al.* The CST Complex Mediates End Protection at Double-Strand Breaks and Promotes PARP Inhibitor Sensitivity in BRCA1-Deficient Cells. *Cell Rep.* **23**, 2107–2118 (2018).
141. Tomida, J. *et al.* FAM 35A associates with REV 7 and modulates DNA damage responses of normal and BRCA 1-defective cells . *EMBO J.* **37**, 1–14 (2018).
142. Noordermeer, S. M. *et al. The shieldin complex mediates 53BP1-dependent DNA repair*. *Nature* vol. 560 (2018).
143. Mirman, Z. *et al.* 53BP1/Rif1/Shieldin counteract DSB resection through CST/ Polα-dependent fill-in. *Nature* **560**, 112–116 (2018).
144. Gupta, R. *et al.* DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell* **173**, 972-988.e23 (2018).
145. Findlay, S. *et al.* SHLD 2/ FAM 35A co-operates with REV 7 to coordinate DNA double-strand break repair pathway choice . *EMBO J.* **37**, 1–20 (2018).
146. Dev, H. *et al.* Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. *Nat. Cell Biol.* **20**, 954–965 (2018).
147. Panier, S. & Boulton, S. J. Double-strand break repair: 53BP1 comes into focus. *Nat. Rev. Mol. Cell Biol.* **15**, 7–18 (2013).
148. Zimmermann, M. & De Lange, T. 53BP1: Pro choice in DNA repair. *Trends Cell Biol.* **24**, 108–117 (2014).
149. Boersma, V. *et al.* MAD2L2 controls DNA repair at telomeres and DNA breaks by inhibiting 5′ end resection. *Nature* **521**, 537–540 (2015).

150. Xu, G. *et al.* REV7 counteracts DNA double-strand break resection and affects PARP inhibition. *Nature* **521**, 541–544 (2015).
151. Buonomo, S. B. C., Wu, Y., Ferguson, D. & De Lange, T. Mammalian Rif1 contributes to replication stress survival and homology-directed repair. *J. Cell Biol.* **187**, 385–398 (2009).
152. Smith, J. S., Caputo, E. & Boeke, J. D. A Genetic Screen for Ribosomal DNA Silencing Defects Identifies Multiple DNA Replication and Chromatin-Modulating Factors. *Mol. Cell. Biol.* **19**, 3184–3197 (1999).
153. Li, P. *et al.* Rif1 promotes a repressive chromatin state to safeguard against endogenous retrovirus activation. *Nucleic Acids Res.* **45**, 12723–12738 (2017).
154. Hardy, C. F. J., Sussel, L. & Shore, D. A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes Dev.* **6**, 801–814 (1992).
155. Dan, J. *et al.* Rif1 Maintains Telomere Length Homeostasis of ESCs by Mediating Heterochromatin Silencing. *Dev. Cell* **29**, 7–19 (2014).
156. Cornacchia, D. *et al.* Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J.* **31**, 3678–3690 (2012).
157. Yamazaki, S. *et al.* Rif1 regulates the replication timing domains on the human genome. *EMBO J.* **31**, 3667–3677 (2012).
158. Xu, D. *et al.* Rif1 provides a new DNA-binding interface for the Bloom syndrome complex to maintain normal replication. *EMBO J.* **29**, 3140–3155 (2010).
159. Hengeveld, R. C. C. *et al.* Rif1 Is Required for Resolution of Ultrafine DNA Bridges in Anaphase to Ensure Genomic Stability. *Dev. Cell* **34**, 466–474 (2015).
160. Silverman, J., Takai, H., Buonomo, S. B. C., Eisenhaber, F. & Lange, T. De. Human Rif1, ortholog of a yeast telomeric protein, is regulated by ATM and 53BPP1 and functions in the S-phase checkpoint. *Genes Dev.* **18**, 2108–2119 (2004).
161. de Villartay, J.-P., Fischer, A. & Durandy, A. The mechanisms of immune diversification and their disorders. *Nat. Rev. Immunol.* **3**, 962–972 (2003).
162. Robbiani, D. F. & Nussenzweig, M. C. Chromosome Translocation, B Cell Lymphoma, and Activation-Induced Cytidine Deaminase. *Annu. Rev. Pathol. Mech. Dis.* **8**, 79–103 (2013).
163. Nussenzweig, A. & Nussenzweig, M. C. Origin of Chromosomal Translocations in Lymphoid Cancer. *Cell* **141**, 27–38 (2010).
164. Alt, F. W., Zhang, Y., Meng, F. L., Guo, C. & Schwer, B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* **152**, 417–429 (2013).
165. Durandy, A., Kracker, S. & Fischer, A. Primary antibody deficiencies. *Nat. Rev. Immunol.* **13**, 519–533 (2013).
166. Fried, A. J. & Bonilla, F. A. Pathogenesis, diagnosis, and management of primary antibody deficiencies and infections. *Clin. Microbiol. Rev.* **22**, 396–414 (2009).
167. Durandy, A., Kracker, S. & Fischer, A. Primary antibody deficiencies. *Nat. Rev. Immunol.* **13**, 519–533 (2013).

168. Zan, H. & Casali, P. Regulation of Aicda expression and AID activity. *Autoimmunity* **46**, 83–101 (2013).
169. Wu, X., Darce, J. R., Chang, S. K., Nowakowski, G. S. & Jelinek, D. F. Alternative splicing regulates activation-induced cytidine deaminase (AID): Implications for suppression of AID mutagenic activity in normal and malignant B cells. *Blood* **112**, 4675–4682 (2008).
170. Dorsett, Y. *et al.* MicroRNA-155 Suppresses Activation-Induced Cytidine Deaminase-Mediated Myc-Igh Translocation. *Immunity* **28**, 630–638 (2008).
171. Teng, G. *et al.* microRNA-155 is a negative regulator of Activation Induced Cytidine deaminase. *Immunity* **28**, 621–629 (2008).
172. Albesiano, E. *et al. Activation-induced cytidine deaminase in chronic lymphocytic leukemia B cells: Expression as multiple forms in a dynamic, variably sized fraction of the clone*. *Blood* vol. 102 (2003).
173. Iacobucci, I. *et al.* Different isoforms of the B-cell mutator activation-induced cytidine deaminase are aberrantly expressed in BCR-ABL1-positive acute lymphoblastic leukemia patients. *Leukemia* **24**, 66–73 (2010).
174. van Maldegem, F., Jibodh, R. A., van Dijk, R., Bende, R. J. & van Noesel, C. J. M. Activation-Induced Cytidine Deaminase Splice Variants Are Defective Because of the Lack of Structural Support for the Catalytic Site. *J. Immunol.* **184**, 2487–2491 (2010).
175. van Maldegem, F. *et al.* AID splice variants lack deaminase activity With. *Blood* **113**, 1862–3 (2009).
176. Gazumyan, A. *et al.* Amino-Terminal Phosphorylation of Activation-Induced Cytidine Deaminase Suppresses c-myc/IgH Translocation. *Mol. Cell. Biol.* **31**, 442–449 (2011).
177. Mu, Y., Zelazowska, M. A. & McBride, K. M. Phosphorylation promotes activation-induced cytidine deaminase activity at the Myc oncogene. *J. Exp. Med.* **214**, 3543–3552 (2017).
178. Ito, S. *et al.* Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl. Acad. Sci.* **101**, 1975–1980 (2004).
179. Brar, S. S., Watson, M. & Diaz, M. Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. *J. Biol. Chem.* **279**, 26395–26401 (2004).
180. Di Noia, J. M. & Patenaude, A.-M. The mechanisms regulating the subcellular localization of AID. *Nucleus* **14**, 325–331 (2010).
181. Küppers, R. Mechanisms of B-cell lymphoma pathogenesis. *Nat. Rev. Cancer* **5**, 251–262 (2005).
182. Franco, S. *et al.* DNA-PKcs and Artemis function in the end-joining phase of immunoglobulin heavy chain class switch recombination. *J. Exp. Med.* **205**, 557–564 (2008).
183. Boboila, C. *et al.* Alternative end-joining catalyzes robust IgH locus deletions and

translocations in the combined absence of ligase 4 and Ku70. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3034–3039 (2010).

184. Wang, J. H. *et al.* Oncogenic transformation in the absence of Xrcc4 targets peripheral B cells that have undergone editing and switching. *J. Exp. Med.* **205**, 3079–3090 (2008).
185. Klemm, L. *et al.* The B Cell Mutator AID Promotes B Lymphoid Blast Crisis and Drug Resistance in Chronic Myeloid Leukemia. *Cancer Cell* **16**, 232–245 (2009).
186. Matsumoto, Y. *et al.* Up-regulation of activation-induced cytidine deaminase causes genetic aberrations at the CDKN2b-CDKN2a in gastric cancer. *Gastroenterology* **139**, 1984–1994 (2010).
187. Marusawa, H. & Chiba, T. Helicobacter pylori-induced activation-induced cytidine deaminase expression and carcinogenesis. *Curr. Opin. Immunol.* **22**, 442–7 (2010).
188. Rickert, R. C., Roes, J. & Rajewsky, K. B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic Acids Res.* **25**, 1317–1318 (1997).
189. Nakamura, M. *et al.* High frequency class switching of an IgM + B lymphoma clone CH12F3 to IgA + cells. *Int. Immunol.* **8**, 193–201 (1996).
190. Han, L. & Yu, K. Altered kinetics of nonhomologous end joining and class switch recombination in ligase IV-deficient B cells. *J. Exp. Med.* **205**, 2745–2753 (2008).
191. Pear, W. S., Nolan, G. P., Scott, M. L. & Baltimore, D. Production of high-titer helper-free retroviruses by transient transfection. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 8392–8396 (1993).
192. Bunting, S. F. *et al.* 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* **141**, 243–254 (2010).
193. Christie, D. A. *et al.* PU.1 Opposes IL-7–Dependent Proliferation of Developing B Cells with Involvement of the Direct Target Gene Bruton Tyrosine Kinase . *J. Immunol.* **194**, 595–605 (2015).
194. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
195. Chu, V. T. *et al.* Efficient CRISPR-mediated mutagenesis in primary immune cells using CrispRGold and a C57BL/6 Cas9 transgenic mouse line. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12514–12519 (2016).
196. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science (80-.).* **343**, 84–87 (2014).
197. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–308 (2013).
198. Ramachandran, S. *et al.* The SAGA Deubiquitination Module Promotes DNA Repair and Class Switch Recombination through ATM and DNAPK-Mediated γH2AX Formation. *Cell Rep.* **15**, 1554–1565 (2016).
199. Takahashi, Y., Ohta, H. & Takemori, T. Fas is required for clonal selection in germinal centers and the subsequent establishment of the memory B cell repertoire. *Immunity* **14**, 181–192 (2001).

200. Ye, S. K. *et al.* Induction of germline transcription in the TCRγ, locus by Stat5: Implications for accessibility control by the IL-7 receptor. *Immunity* **11**, 213–223 (1999).
201. Sander, S. *et al.* PI3 Kinase and FOXO1 Transcription Factor Activity Differentially Control B Cells in the Germinal Center Light and Dark Zones. *Immunity* **43**, 1075–1086 (2015).
202. Oruc, Z., Boumédiène, A., Le Bert, M. & Khamlichi, A. A. Replacement of Iγ3 germ-line promoter by Iγ1 inhibits class-switch recombination to IgG3. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 20484–20489 (2007).
203. Xu, G. *et al.* REV7 counteracts DNA double-strand break resection and affects PARP inhibition. *Nature* **advance on**, (2015).
204. Bray, N. L., Pimentel, H., Melsted, P. & Patcher, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–7 (2016).
205. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences [version 2; referees: 2 approved]. *F1000Research* **4**, 1–22 (2016).
206. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
207. Wurmus, R. *et al.* PiGx: Reproducible genomics analysis pipelines with GNU Guix. *Gigascience* **7**, 1–14 (2018).
208. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
209. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
210. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
211. Subbotin, R. I. & Chait, B. T. A pipeline for determining protein-protein interactions and proximities in the cellular milieu. *Mol. Cell. Proteomics* **13**, 2824–2835 (2014).
212. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
213. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
214. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**, 576–589 (2010).
215. Yu, G., Wang, L. G. & He, Q. Y. ChIP seeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
216. Bishop, G. a & Haughton, G. Induced differentiation of a transformed clone of Ly-1+ B cells by clonal T cells and antigen. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 7410–7414 (1986).

217. Le Cong *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science (80-.).* **339**, 819–824 (2013).
218. Bothmer, A., Robbiani, D. & Virgilio, M. Di. Regulation of DNA end joining, resection, and immunoglobulin class switch recombination by 53BP1. *Mol. Cell* **42**, 319–329 (2011).
219. Ward, I. M. *et al.* 53BP1 is required for class switch recombination. *J. Cell Biol.* **165**, 459–464 (2004).
220. Tackett, A. *et al.* I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. *J. Proteome Res.* **4**, 1752–56 (2005).
221. Subbotin, R. I. & Chait, B. T. A Pipeline for Determining Protein-Protein Interactions and Proximities in the Cellular Milieu. *Mol. Cell. Proteomics* 2824–2835 (2014) doi:10.1074/mcp.M114.041095.
222. Sabatini, D. D., Bensch, K. & Barnett, R. J. Cytochemistry and electron microscopy. The preservation of cellular ultrastructure and enzymatic activity by aldehyde fixation. *J. Cell Biol.* **17**, 19–58 (1963).
223. Walt, D. R. & Agayn, V. I. The chemistry of enzyme and protein immobilization with glutaraldehyde. *Trends Anal. Chem.* **13**, 425–30 (1994).
224. Käll, L., Storey, J. D., MacCoss, M. J. & Noble, W. S. Posterior error probabilities and false discovery rates: Two sides of the same coin. *J. Proteome Res.* **7**, 40–44 (2008).
225. Adhikary, S. *et al.* Selective recognition of H3.1K36 dimethylation/H4K16 acetylation facilitates the regulation of all-trans-retinoic acid (ATRA)-responsive genes by putative chromatin reader ZMYND8. *J. Biol. Chem.* **291**, 2664–2681 (2016).
226. Ghosh, K. *et al.* Positive Regulation of Transcription by Human ZMYND8 through Its Association with P-TEFb Complex. *Cell Rep.* **24**, 2141-2154.e6 (2018).
227. Shen, H. *et al.* Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase Complex. *Cell* **165**, 331–342 (2016).
228. Li, N. *et al.* ZMYND8 Reads the Dual Histone Mark H3K4me1-H3K14ac to Antagonize the Expression of Metastasis-Linked Genes. *Mol. Cell* **63**, 470–484 (2016).
229. Gong, F. *et al.* Screen identifies bromodomain protein ZMYND8 in chromatin recognition of transcription-associated DNA damage that promotes homologous recombination. *Genes Dev.* **29**, 197–211 (2015).
230. Gong, F., Clouaire, T., Aguirrebengoa, M., Legube, G. & Miller, K. M. Histone demethylase KDM5A regulates the ZMY ND8-NuRD chromatin remodeler to promote DNA repair. *J. Cell Biol.* **216**, 1959–1974 (2017).
231. Kloet, S. L. *et al.* Towards elucidating the stability, dynamics and architecture of the nucleosome remodeling and deacetylase complex by using quantitative interaction proteomics. *FEBS J.* **282**, 1774–1785 (2015).
232. Spruijt, C. G. *et al.* ZMYND8 Co-localizes with NuRD on Target Genes and Regulates Poly(ADP-Ribose)-Dependent Recruitment of GATAD2A/NuRD to Sites of DNA Damage. *Cell Rep.* **17**, 783–798 (2016).
233. Xia, L. *et al.* CHD4 Has Oncogenic Functions in Initiating and Maintaining Epigenetic

Suppression of Multiple Tumor Suppressor Genes. *Cancer Cell* **31**, 653-668.e7 (2017).

234. Marygold, S. J. *et al.* The ribosomal protein genes and Minute loci of Drosophila melanogaster. *Genome Biol.* **8**, (2007).
235. Brown, A., Amunts, A., Bai, X., Sugimoto, Y. & Edwards, P. C. Structure of the large ribosomal subunit from human mitochondria. *Science (80-.).* **346**, 718–722 (2014).
236. Fung, S., Nishimura, T., Sasarman, F. & Shoubridge, E. A. The conserved interaction of C7orf30 with MRPL14 promotes biogenesis of the mitochondrial large ribosomal subunit and mitochondrial translation. *Mol. Biol. Cell* **24**, 184–193 (2013).
237. Imataka, H., Olsen, H. S. & Sonenberg, N. A new translational regulator with homology to eukaryotic translation initiation factor 4G. *EMBO J.* **16**, 817–825 (1997).
238. Henis-Korenblit, S. *et al.* The caspase-cleaved DAP5 protein supports internal ribosome entry site-mediated translation of death proteins. *Proc. Natl. Acad. Sci.* **99**, 5400–5405 (2002).
239. Marash, L. *et al.* DAP5 Promotes Cap-Independent Translation of Bcl-2 and CDK1 to Facilitate Cell Survival during Mitosis. *Mol. Cell* **30**, 447–459 (2008).
240. Shen, L., Huang, K., Chen, H. & Huang, F. L. Molecular Cloning and Characterization of a Novel Casein Kinase II Substrate, HASPP28, from Rat Brain. *Arch. Biochem. Biophys.* **327**, 131–141 (1996).
241. Fischer, W. H. & Schubert, D. Characterization of a novel platelet-derived growth factor-associated protein. *J. Neurochem.* **66**, 2213–2216 (1996).
242. Trendel, J. *et al.* The Human RNA-Binding Proteome and Its Dynamics During Arsenite-Induced Translational Arrest. *Cell* **176**, 391-403.e19 (2018).
243. Baltz, A. G. *et al.* The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Mol. Cell* **46**, 674–690 (2012).
244. Castello, A. *et al.* Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* **149**, 1393–1406 (2012).
245. Iadevaia, V. *et al.* Tandem RNA isolation reveals functional rearrangement of RNA-binding proteins on CDKN1B/p27Kip1 3'UTRs in cisplatin treated cells. *RNA Biol.* **17**, 33–46 (2019).
246. Castello, A. *et al.* Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell* **63**, 696–710 (2016).
247. Sharma, V. *et al.* Increased expression of platelet-derived growth factor associated protein-1 is associated with PDGF-B mediated glioma progression. *Int J Biochem Cell Biol.* **78**, 194–205 (2016).
248. Weston, V. J., Wei, W., Stankovic, T. & Kearns, P. Synergistic action of dual IGF1/R and MEK inhibition sensitizes childhood acute lymphoblastic leukemia (ALL) cells to cytotoxic agents and involves downregulation of STAT6 and PDAP1. *Exp. Hematol.* **63**, 52-63.e5 (2018).
249. Maizels, N. & Bothwell, A. The T-cell-independent immune response to the hapten NP uses a large repertoire of heavy chain genes. *Cell* **43**, 715–720 (1985).

250. Quah, B. & Parish, C. New and improved methods for measuring lymphocyte proliferation in vitro and in vivo using CFSE-like fluorescent dyes. *J. Immunol. Methods* **379**, 1–14 (2012).
251. Satoh, M. & Lindahl, T. Role of poly(ADP-ribose) formation in DNA repair. *Nature* **356**, 356–58 (1992).
252. Underhill, C., Toulmonde, M. & Bonnefoi, H. A review of PARP inhibitors: From bench to bedside. *Ann. Oncol.* **22**, 268–279 (2011).
253. Helleday, T. The underlying mechanism for the PARP and BRCA synthetic lethality: Clearing up the misunderstandings. *Mol. Oncol.* **5**, 387–393 (2011).
254. Le, M. X. *et al.* Kin17 facilitates multiple double-strand break repair pathways that govern B cell class switching. *Sci. Rep.* **6**, 1–13 (2016).
255. Sernández, I. V., de Yébenes, V. G., Dorsett, Y. & Ramiro, A. R. Haploinsufficiency of activation-induced deaminase for antibody diversification and chromosome translocations both in vitro and in vivo. *PLoS One* **3**, 1–9 (2008).
256. Takizawa, M. *et al.* AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *J. Exp. Med.* **205**, 1949–1957 (2008).
257. Robbiani, D. F. *et al.* Plasmodium Infection Promotes Genomic Instability and AID-Dependent B Cell Lymphoma. *Cell* **162**, 727–737 (2015).
258. Jonkers, I. & Lis, J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 167–177 (2015).
259. Jolly, C. J., Klix, N. & Neuberger, M. S. Rapid methods for the analysis of immunoglobulin gene hypermutation: Application to transgenic and gene targeted mice. *Nucleic Acids Res.* **25**, 1913–1919 (1997).
260. González-Fernández, Á., Gilmore, D. & Milstein, C. Age-related decrease in the proportion of germinal center B cells from mouse Peyer's patches is accompanied by an accumulation of somatic mutations in their immunoglobulin genes. *Eur. J. Immunol.* **24**, 2918–21 (1994).
261. Reboldi, A. & Cyster, J. G. Peyer's patches: Organizing B cell responses at the intestinal frontier. *Immunol. Rev.* **271**, 230–245 (2016).
262. Schlissel, M. & Baltimore, D. Activation of immunoglobulin kappa gene rearrangement correlates with induction of germline kappa gene transcription. *Cell* **9**, 1001–7 (1989).
263. Xiang, Y. & Garrard, W. T. The Downstream Transcriptional Enhancer, Ed, Positively Regulates Mouse Igκ Gene Expression and Somatic Hypermutation. *J. Immunol.* **180**, 6725–6732 (2008).
264. Chen, C. Y. A., Ezzeddine, N. & Shyu, A. Bin. Messenger RNA Half-Life Measurements in Mammalian Cells. *Methods Enzymol.* **448**, 335–357 (2008).
265. Fowler, T. *et al.* Divergence of transcriptional landscape occurs early in B cell activation. *Epigenetics and Chromatin* **8**, 1–14 (2015).
266. Sundaravinayagam, D. *et al.* 53BP1 Supports Immunoglobulin Class Switch

Recombination Independently of Its DNA Double-Strand Break End Protection Function. *Cell Rep.* **28**, 1389-1399.e6 (2019).

267. Tanaka, H. *et al.* Epigenetic Regulation of the Blimp-1 Gene (Prdm1) in B Cells Involves Bach2 and Histone Deacetylase 3. *J. Biol. Chem.* **291**, 6316–6330 (2016).
268. Loh, Y. H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431–440 (2006).
269. Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. *Cell* **132**, 1049–1061 (2008).
270. Daxinger, L. *et al.* An ENU mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. *Genome Biol.* **14**, (2013).
271. Wang, J. *et al.* A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364–368 (2006).
272. Raviram, R., Rocha, P. P., Bonneau, R. & Skok, J. A. Interpreting 4C-Seq data: How far can we go? *Epigenomics* **6**, 455–457 (2014).
273. Verma-Gaur, J. *et al.* Noncoding transcription within the Igh distal VH region at PAIR elements affects the 3D structure of the Igh locus in pro-B cells. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17004–17009 (2012).
274. Ni, Z. *et al.* P-TEFb Is Critical for the Maturation of RNA Polymerase II into Productive Elongation In Vivo. *Mol. Cell. Biol.* **28**, 1161–1170 (2008).
275. Yu, Y. *et al.* TRIB3 mediates the expression of Wnt5a and activation of nucelar factor-kB in Porphyromonas endodontalis lipopolysaccharide-treated osteoblasts. *Mol Oral Microbiol.* **30**, 295–306 (2015).
276. Wu, M., Xu, L. G., Zhai, Z. & Shu, H. B. SINK is a p65-interacting negative regulator of NF-κB-dependent transcription. *J. Biol. Chem.* **278**, 27072–27079 (2003).
277. Endo, Y. *et al.* Expression of activation-induced cytidine deaminase in human hepatocytes via NF-κB signaling. *Oncogene* **26**, 5587–5595 (2007).
278. Park, S. *et al.* APRIL stimulates NF-kB-mediated HoxC4 induction for AID expression in mouse B cells. *Cytokine* **61**, 608–13 (2013).
279. Ise, W. *et al.* BATF controls the global regulators of CSR in both B and T cells. *Nat. Immunol.* **12**, 536–543 (2011).
280. Santos, J. M. *et al.* Duplication of a germline promoter downstream of the IgH 3′ regulatory region impairs class switch recombination. *Sci. Rep.* **8**, 1–6 (2018).
281. Fear, D. J., McCloskey, N., O'Connor, B., Felsenfeld, G. & Gould, H. J. Transcription of Ig Germline Genes in Single Human B Cells and the Role of Cytokines in Isotype Determination. *J. Immunol.* **173**, 4529–4538 (2004).
282. Pakos-Zebrucka, K. *et al.* The integrated stress response. *EMBO Rep.* **17**, 1374–1395 (2016).
283. Walter, P. & Ron, D. The unfolded protein response: from stress pathway to homeostatic regulation. *Science (80-.).* **34**, 1081–86 (2011).
284. Harding, H. P., Zhang, Y., Bertolotti, A., Zeng, H. & Ron, D. Perk is essential for

translational regulation and cell survival during the unfolded protein response. *Mol. Cell* **5**, 897–904 (2000).

285. Harding, H. P. *et al.* An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol. Cell* **11**, 619–633 (2003).
286. Brewer, J. & Hendershot, L. Building an antibody factory: a job of the unfolded protein response. *Nat. Immunol.* **6**, 23–9 (2005).
287. Calame, K., Lin, K. & Tunyaplin, C. Regulatory mechanisms that determine the development and function of plasma cells. *Annu. Rev. Immunol.* **21**, 205–30 (2003).
288. Wiest, D. L. *et al.* Membrane biogenesis during B cell differentiation: Most endoplasmic reticulum proteins are expressed coordinately. *J. Cell Biol.* **110**, 1501–1511 (1990).
289. Manuscript, A. The unfolded protein response of B-lymphocytes: PERK-independent development of antibody-secreting cells. *Gass, JN. Jiang, HY. Wek, RC. Brew. JW.* **45**, 1035–43 (2008).
290. Shaffer, A. L. *et al.* XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* **21**, 81–93 (2004).
291. Todd, D. J. *et al.* XBP1 governs late events in plasma cell differentiation and is not required for antigen-specific memory B cell development. *J. Exp. Med.* **206**, 2151–2159 (2009).
292. Boothby, M. & Rickert, R. C. Metabolic Regulation of the Immune Humoral Response. *Immunity* **46**, 743–755 (2017).
293. Dufort, F. J. *et al.* Cutting Edge: IL-4-Mediated Protection of Primary B Lymphocytes from Apoptosis via Stat6-Dependent Regulation of Glycolytic Metabolism. *J. Immunol.* **179**, 4953–4957 (2007).
294. Doughty, C. A. *et al.* Antigen receptor-mediated changes in glucose metabolism in B lymphocytes: Role of phosphatidylinositol 3-kinase signaling in the glycolytic control of growth. *Blood* **107**, 4458–4465 (2006).
295. Caro-maldonado, A., Wang, R., Nichols, A. G. & Kuraoka, M. Metabolic Reprogramming is Required for Antibody Production That is Suppressed in Anergic but Exaggerated in Chronically BAFF-Exposed B cells. *J Immunol* **192**, 3626–3636 (2014).
296. Waters, L. R., Ahsan, F. M., Wolf, D. M., Shirihai, O. & Teitell, M. A. Initial B Cell Activation Induces Metabolic Reprogramming and Mitochondrial Remodeling. *iScience* **5**, 99–109 (2018).
297. Jellusova, J. & Rickert, R. A brake for B cell proliferation: appropiate responses to metabolic stress are crucial to maintain B cell viability and prevent malignant outgrowth. *BioEssays* **39**, (2017).
298. Jang, K. J. *et al.* Mitochondrial function provides instructive signals for activation-induced B-cell fates. *Nat. Commun.* **6**, (2015).
299. Quirós, P. M. *et al.* Multi-omics analysis identifies ATF4 as a key regulator of the mitochondrial stress response in mammals. *J. Cell Biol.* **216**, 2027–2045 (2017).

300. Yang, X. *et al.* ATF4 Regulates CD4 + T Cell Immune Responses through Metabolic Reprogramming. *Cell Rep.* **23**, 1754–1766 (2018).
301. Ohoka, N., Yoshii, S., Hattori, T., Onozaki, K. & Hayashi, H. TRB3, a novel ER stress-inducible gene, is induced via ATF4-CHOP pathway and is involved in cell death. *EMBO J.* **24**, 1243–1255 (2005).
302. Qu, J. *et al.* TRIB3 suppresses proliferation and invasion and promotes apoptosis of endometrial cancer cells by regulating the AKT signaling pathway. *Onco. Targets. Ther.* **12**, 2235–2245 (2019).
303. Rozpedek, W. *et al.* The Role of the PERK/eIF2α/ATF4/CHOP Signaling Pathway in Tumor Progression During Endoplasmic Reticulum Stress. *Curr. Mol. Med.* **16**, 533–544 (2016).
304. Hu, H., Tian, M., Ding, C. & Yu, S. The C/EBP homologous protein (CHOP) transcription factor functions in endoplasmic reticulum stress-induced apoptosis and microbial infection. *Front. Immunol.* **10**, 1–13 (2019).
305. Albershardt, T. C., Iritani, B. M. & Ruddell, A. Evaluation of reference genes for quantitative PCR analysis of mouse lymphocytes. *J. Immunol. Methods* **348**, 196–199 (2012).

ABBREVIATIONS

3'E_κ	3' kappa enhancer
3'RR	3' regulatory region
4-HT	4-hydroxytamoxifen
4C-se	circularized chromosome conformation capture combined with sequencing
53BP1	p53-binding protein 1
A	adenine
ACK	ammonium-chloride-potassium
AID	activation-induced deaminase
alt-NHEJ	alternative non-homologous end-joining
APE1/2	apurinic/apyrimidinic endonuclease 1/2
ATM	ataxia telangiectasia mutated
BAFF	B-cell activator of the TNF-α family
BAFF-R	BAFF-receptor
BCR	B cell receptor
BER	base excision repair
BLIMP-1	B-lymphocyte-induced maturation protein-1
BM	bone marrow
BRCA1	breast cancer 1
BSA	bovine serum albumin
C	cytosine
c-NHEJ	classical non-homologous end-joining
CD40L	CD40 ligand
CFSE	carboxyfluorescein succinimidyl ester
ChIP-seq	chromatin immunoprecipitation combined with sequencing
CLP	common lymphoid progenitor
CMP	common myeloid progenitor
CRISPR	clustered regularly interspaced short palindromic repeats
CSR	class switch recombination
CtIP	C-terminal-binding protein interacting protein
Ctrl	control
CTV	cell trace violet
DDR	DNA damage response
DMEM	Dulbecco's modified eagle medium
DNA-PKcs	DNA-PK catalytic subunit

DSB	double strand break
ECAR	extracellular acidification rate
EDTA	ethylenediaminetetraacetic acid
EIF4G2	eukaryotic translation initiation factor 4 gamma 2
Ei_{κ}	intronic kappa enhancer
ELISA	enzyme-linked immunosorbent assay
ELISpot	enzyme-linked immunospot
ER	endoplasmic reticulum
eRNA	enhancer RNA
EXO-1	exonuclease-1
$E_{\lambda}2\text{-}4$	lambda 2-4 enhancer
$E_{\lambda}3\text{-}1$	lambda 3-1 enhancer
Eμ	enhancer μ region
Fab	antigen binding fragment
FBS	fetal bovine serum
Fc	crystallizable fragment
FDR	false discovery rate
FL	full-length
FO	follicular
G	guanine
GC	germinal center
gDNA	genomic DNA
GFP	green fluorescent protein
GLT	germline transcription
gRNA	guide RNA
GRO-seq	global run-on sequencing
H2AX	H2A histone family member X
HA	hemagglutinin
HEPES	4-(2-hydroxylethyl)-1-piperazineethanesulfonic acid
HP1γ	heterochromatin protein 1
HR	homologous recombination
HRP	horseradish peroxidase
hs	DNase hypersensitive sites
HSC	hematopoietic stem cell
I-DIRT	isotopic differentiation of interactions as random or targeted
IFN-γ	interferon-γ
Ig	immunoglobulin

IgH	immunoglobulin heavy chain
IgL	immunoglobulin light chain
IL-4	interleukin-4
IL-5	interleukin-5
iMEFs	immortalized mouse embryonic fibroblasts
IR	ionized radiation
IRES	internal ribosome entry site
ISR	integrated stress response
KAP1	KRAB domain-associated protein 1
KO	knock-out
LIG4	DNA ligase 4
LLPC	long-lived plasma cell
Log2FC	Log 2 fold change
LPS	lipopolysaccharide
MAD2L2	mitotic arrest deficient-like 2
MDC1	mediator of DNA damage checkpoint 1
MLL	mixed-lineage leukemia-like
MMP	multipotent progenitor
MMR	mismatch repair pathway
MRE11	meiotic recombination 11 homolog A
mRNA	messenger RNA
MSH2/MSH6	homologue of bacterial MutS 2/6
MZ	marginal zone
NBS1	Nijmegen breakage syndrome 1
NES	nuclear export signal
NF-κB	nuclear factor kappa beta
NK	natural killer
NLS	nuclear localization signal
NP-CGG	4-Hydroxy-3-nitrophenylacetyl hapten conjugated to chicken gamma globulin
OCR	oxygen consumption rate
OXPHOS	oxidative phosphorylation
P-TEFb	positive transcription elongation factor b
PAD	primary antibody deficiency
PAMPs	pathogen-associated molecular patterns
PARC-CLIP	photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation

PARPi	PARP inhibitor
PB	plasmablast
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDAP1	platelet-derived growth factor-associated protein 1
pre-B	precursor B cell
pre-BCR	pre-B cell receptor
pre-pro B cell	pre-progenitor B cell
pro-B	progenitor B cell
PTIP	pax interaction with transcription-activation domain protein 1
RAG	recombination activating gene
RIF1	replication timing regulatory factor 1
RMP	reads per gene per million mapped sequence reads
RNAPII	RNA polymerase II
RNAPII-S2	RNA polymerase II – phosphorylated on serine 2
RNAPII-S5	RNA polymerase II – phosphorylated on serine 5
RNF168	ring finger protein 168
RNF8	ring finger protein 8
RPMI	Roswell Park Memorial Institute
RS	recombining segment
RSS	recombination signal sequence
RT	room temperature
RT-qPCR	quantitative reverse transcription PCR
S3	serine-3
S38	serine-38
SDS	sodium dodecyl sulfate
SHM	somatic hypermutation
SILAC	stable isotope labeling with amino acids in cell culture
SLC	surrogate light chain
SLPC	short-lived plasma cell
SPF	specific pathogen-free
SPT5	suppressor of Ty 5 homolog
SSB	single strand break
ssDNA	single-stranded DNA
T	timine
T1, T2 and T3	Transitional B cell populations 1, 2 and 3
T140	threonine-140

TCA	tricaboxylic acid cycle
TD	T cell-dependent
TGFβ	transforming growth factor β
TI	T cell-independent
TLR	toll-like receptors
TSS	transcription start sties
UDR	ubiquitination-dependent recruitment
UNG	uracil DNA-glycosilase
UPR	unfolded protein response
V(D)J	variable, (diversity), joining
WB	Western Blot
WCE	whole-cell-extracts
WT	wild-type
XLF	XRCC4-like factor
XRCC4	X-ray repair cross complementing 4
ZMYND8	zinc finger MYND-type (myeloid, nervy, and DEAF-1) containing 8
K	Kappa
λ	Lambda

STATEMENT OF CONTRIBUTION

The ZMYND8 project has been conducted in collaboration with Daniel B. Rosen (The Rockefeller University, NY). All results related to this project have been included for completeness, but individual contributions are indicated in this list:

Figure 4.2. Screening for novel RIF1-interacting CSR factors in B-lymphocytes.
Panel A and B: Daniel B. Rosen, in collaboration with Brian T. Chait (RU)
Panel C: gRNAs designed by V. Delgado-Benito and cloned by L. Keller. Screen performed by D.B. Rosen and V. Delgado-Benito.

Figure 4.4. Deletion of *Zmynd8* in CH12 causes loss of CSR.
Panel B: D.B. Rosen and V. Delgado-Benito.
Panel C: experiment performed by V. Delgado-Benito, with different ZMYND8-KO clones generated by D.B. Rosen and V. Delgado-Benito. WB performed by D.B. Rosen.

Figure 4.5. Generation of a conditional *Zmynd8* knock-out mouse model
V. Delgado-Benito.

Figure 4.6. ZMYND8 is dispensable for B cell development.
V. Delgado-Benito.

Figure 4.7. ZMYND8 is a novel factor required for CSR *ex vivo*.
V. Delgado-Benito.

Figure 4.8. ZMYND8 is a novel factor required for CSR *in vivo*.
V. Delgado-Benito.

Figure 4.9. ZMYND8 is dispensable for B cell proliferation.
V. Delgado-Benito.

Figure 4.10. Generation of *Zmynd8-KO* iMEFs.
V. Delgado-Benito.

Figure 4.11. ZMYND8 is dispensable for the signaling of stochastic DNA damage.
V. Delgado-Benito and D. Sundaravinayagam.

Figure 4.12. ZMYND8 is dispensable for the repair of Cas9-induced DSBs.
V. Delgado-Benito.

Figure 4.13. ZMYND8 is dispensable for the repair of IR- and PARPi-induced DSBs.
Panels A-B: V. Delgado-Benito and D. Sundaravinayagam.
Panel C: V. Delgado-Benito.

Figure 4.14. ZMYND8 is dispensable for AID expression and induction of DSBs.
Panels A: V. Delgado-Benito
Panel B: V. Delgado-Benito performed the experiments. J.A. Pai and T.Y. Oliveira conducted the computational analysis.

Figure 4.15. ZMYND8 is dispensable for Sμ (donor) GLT.
V. Delgado-Benito.

Figure 4.16. ZMYND8 supports GLT of acceptor S regions.
V. Delgado-Benito.

Figure 4.17. ZMYND8 binds to *Igh* enhancers and regulates RNAPII loading at the 3'RR.
V. Delgado-Benito isolated, cultured and fixed the splenocytes; D.B. Rosen prepared the library for ChIP-seq; JA. Pai and TY. Oliveira conducted the computational analysis.

Figure 4.18. ZMYND8 represses hs1,2 and hs3b transcription.
V. Delgado-Benito.

Figure 4.19. ZMYND8 is required for efficient SHM at the *Igh* locus.
V. Delgado-Benito.

Figure S1. Sμ-Sγ1 junction recombination analysis in *Zmynd8*$^{F/F}$*Cd19*$^{Cre/+}$ and control B cells.
V. Delgado-Benito and M. Driesner.

Figure S2. RNA-seq analysis of ZMYND8-deficient CH12 cells.
D.B. Rosen prepared the library for ChIP-seq; JA. Pai and TY. Oliveira conducted the computational analysis.

Figure S3. RNAPII-S5 is increased at the 3'RR in ZMYND8- but not in RIF1-deficient CH12 cells.
D.B. Rosen and Q. Wang prepared the library for ChIP-seq; JA. Pai and TY. Oliveira conducted the computational analysis.

ACKNOWLEDGMENTS

There are many reasons why I love Science. First, it is a continuous puzzle and intellectual challenge. It is awesome to have an idea, translate it into an experiment and have a result that is a step forward in understanding how a biological process works. And it is even more awesome to realize about the number of things you learn when you fail. Second, Science gives you the opportunity to develop many different professional and personal skills. You do not realize about it until you take a moment to step out and think about it from an outside perspective. And then there is the cherry on top of the cake, which is that Science exposes you to an international environment. It is so cool to exchange knowledge and personal experiences with people that have been raised differently from you.

Doing a PhD has fueled my passion for Science and helped me to develop professionally and personally for five years. It has been a huge dynamic platform to learn, grow up, meet people from all kinds of cultures and visit places. I came to Berlin with the support of my family and friends and I had the chance to meet incredible people that walked this way with me. I feel tremendously grateful and here are just few words that try to express this feeling:

Thank you, Michela, for believing in me since the first moment we met, your endless support and teaching me how to be a scientist. We did a lot of great things together and we had fun in the process, which made for sure a difference.

Thanks to all the people in the lab for creating a great team and a good environment to do Science together. Thank you, Dev, Tannu, María, Vio, Robert, Jasmin, Madlen, Ali, Lisa, Sandhya and Matteo for your continuous help during these years.

Thanks to Michel Nussenzweig and the people from his lab that I collaborated with. It was an honor for me to be part of the ZMYND8 project and participate in all scientific discussions. I learnt a lot.

Thanks to Annette and Michaela from the PhD office and Hans from the FACS facility, who dedicated time to help me every time I needed it.

Thanks to Klaus Rajewsky and his lab, Kathrin de la Rosa, Gaetano Gargiulo, Uta Höpken, and Niccolò Zampieri for all your advices and inspiration.

Thanks to my doctoral committee members, Claus Scheidereit, Ralf Kühn and Jana Wolf for their constructive criticism. Thank you, Achim Leutz, for your time and your support.

Thanks to MA and SB. What an emotional blast. I just cannot imagine a life without you two.
Thanks to my berliner freerider family, I love you all. I am really speechless. If I knew I was going to have such a strong connection with so many people from different backgrounds. If I knew I was coming to Berlin to speak italian instead of german. If I knew we were going to discover a new planet together. If I knew…

Gracias a radiopatio. Sin (pon tu nombre aquí) ninguno seríamos lo mismo y tampoco habríamos batido el récord en reírnos. Ca uno es ca uno y todos juntos somos inmortales del humor y la química (algunos más pafuera y otros más padentro). Solo necesito cerrar los ojos para conectarme a cada uno de vosotros.

Gracias María, Juanfran, Alba e Itzíar por enseñarme a crear hogar fuera de casa por primera vez y por seguir compartiendo aventuras (ojo al 2020, que se viene).

Gracias Asun, por tu apoyo y compartir momentos en Berlín.

Gracias a sin probleeeemas he aprendido que la distancia y el tiempo no cuentan.

Gracias Pati por estar siempre cerca.

Gracias Ana, por ser parte de mi esté donde esté y pase lo que pase.

Gracias Juan. Gracias por hacerme mejor persona día a día y acompañarme en este viaje sideral. Extragalactic fusion. Do you remember the 21st night of September?

Gracias a mis padres. Todo lo que soy y consigo es un reflejo de la unión de dos personas extraordinarias que siempre han hecho que la fuerza me acompañe a ritmo de “Twist and Shout”. Vosotros me habéis enseñado que el amor no tiene límites.

ERKLÄRUNG

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben.
Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Doktorgrad.
Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde.
Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin vom 5. März 2015.
Weiterhin erkläre ich, dass keine Zusammenarbeit mit gewerblichen Promotionsbearbeiterinnen/Promotionsberatern stattgefunden hat und dass die Grundsätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis eingehalten wurden.

Berlin,

PUBLICATIONS

The Chromatin Reader ZMYND8 Regulates *Igh* Enhancers to Promote Immunoglobulin Class Switch Recombination.

Delgado-Benito V.*, Rosen D.B.*, Wang Q., Gazumyan A., Pai J.A., Oliveira TY., Sundaravinayagam D., Zhang W., Andreani M., Keller L., Kieffer-Kwon K.R., Pękowska A., Jung S., Driesner M., Subbotin R.I., Casellas R., Chait B.T., Nussenzweig M.C., and Di Virgilio, M. **Shared first co-authorship*

Mol. Cell. 72-4, 636-649 (2018)

53BP1 supports Immunoglobulin Class Switch Recombination independently of its DNA end protection function.

Sundaravinayagam D., Rahjouei A., Andreani M., Tupina D., Balasubramanian S., Saha T., Delgado-Benito V., Coralluzzo V., Daumke O., and Di Virgilio M.

Cell Reports. 28-6, 1389-1399 (2019)

PDAP1 protects mature B-lymphocytes from stress-induced death and promotes antibody gene diversification.

Delgado-Benito V., Berruezo-Llacuna M., Altwasser R., Winkler W., Balasubramanian S., Sundaravinayagam D., Graf R., Rahjouei A., Driesner M., Keller L., Janz M., Akalin A., and Di Virgilio M.

Submitted manuscript (2019)